Remote Usability Testing of Online Payments with Smartphones

Ismael Sánchez

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 26.09.2021

Thesis supervisor:

D.Sc. Mika P. Nieminen

Thesis advisor:

Ph.D. Cristian Moral

Aalto University
School of Science

AALTO UNIVERSITY
SCHOOL OF SCIENCE

ABSTRACT OF THE
MASTER'S THESIS

Author: Ismael Sánchez

Title: Remote Usability Testing of Online Payments with Smartphones

Date: 26.09.2021 Language: English Number of pages: 4+80

Department of Computer Science

Major: ICT Innovation SCI3020.

Supervisor: D.Sc. Mika P. Nieminen

Advisor: Ph.D. Cristian Moral

The growth of online technologies applied to the economy is leading to a great appearance of payment applications. To guarantee a correct user experience, it is necessary to evaluate the usability of these applications.
Remote usability testing is an established technique, but how to conduct it, especially for evaluating mobile applications, is not well documented or standardized. This project starts from the question of whether it is possible to perform remote usability testing with mobile payment applications. It also seeks to answer what tools are necessary to perform this testing.
To this end, a study is conducted on the current situation of payment applications and their trends. This is followed by a study of the usability testing methodology and existing techniques, as well as evaluation tools.
With this in mind, it was decided to propose a remote usability testing process for mobile applications. For this purpose, an analysis of the existing needs to conduct this study is carried out. Then the available tools and software are analyzed.
After two iterations of experimentation with users, a procedure to conduct remote usability testing with mobile applications is proved as useful.
Some tools and methodology alternatives are also offered to be able to adapt the procedure to different case studies, making it more generalizable.
Finally, the level of satisfaction with the results obtained is analyzed and future lines of research and improvement are presented.

Keywords: Mobile payments, payment application, usability evaluation, usability testing, remote testing, user experience

Contents

1 Introduction

Technological progress continues changing the way we function in society. Our day to day is also changing because of this. These changes also are changing the financial reality. In particular, mobile banking is one of the most important strategic changes to occur in retail banking in more than a decade [1].

Like credit cards before, the appearance of mobile banking and payment applications is leading to a great change in our payment habits. Apart from allowing payments to be made quickly and easily, these applications allow a wide variety of operations that previously had to be carried out in person at a bank, an automated teller machines (ATM) or a store. This has clear benefits for users saving hassle and time. In turn, this also brings benefits for stores and companies, being able to save on staff and reduce queues among others benefits [1].

The number and popularity of payment applications, especially mobile, continues to rise and pushing to surpass the use of cash and credit cards for regular payments [2]. In recent years, there has been an increase of up to 30% in the global use of this technology, reaching 2.1 billion users in 2019 [3]. These numbers are quite significant considering that the popularity of this technology began to spread in 2015 [4].

For all this kind of applications it's necessary to prove their effectiveness to guarantee an ideal user experience. It is also needed to guarantee the correct flow during use. This is important due to the high frequency of use of these apps as substitute for other payment methods.

It should also be mentioned that, like any emerging technology, payments apps can carry new dangers. These dangers can create fear and mistrust, especially in older users [5]. This becomes of special importance when dealing with money and private and critical information.

Performing an effective evaluation of the user experience can brings benefits in terms of user satisfaction and preference. Also, detecting usability flaws in the design in the early stages of development can save time and expenses [6]. In addition, the effect on the brand image and sales when the product is release with poor usability can have a great economic impact.

Remote usability testing is a common usability evaluation technique. Usability tests are usually conducted in a laboratory where the user is guided by an moderator. Remotely, the user does not have to be in the same room as the moderator.

This technique has difficulties both in the human and technical aspects. Catching all the communicative nuances can be harder when the user can only be seen partially trough a webcam. Running the prototype and monitoring user interaction is also more complicated.

Conducting this type of remote testing open multiple new opportunities. There are proven advantages of remote testing over traditional usability tests in a laboratory, such as reduced costs. It also makes it easier to recruit and work with users in different locations [7]. Additionally, remote testing can potentially provide data from larger numbers of users [8]. Due to these benefits, this working method can be of special interest for developing of new payment applications.

This interest has been further highlighted by the exceptional situation of the

COVID-19 global pandemic, which may limit conducting interviews and face-to-face evaluations. As in other fields such as telehealth, mental health applications and e-commerce [9], this difficulty urges the need to innovate in the field of Human-Computer Interaction. This should be done in order to adapt to other possible situations such as the COVID pandemic and take advantage of the benefits of remote usability testing.

The main objective of usability testing is to find flaws in the design. Despite the benefits of remote usability testing, it is important to guarantee this finding of flaws. With an appropriate and careful process, results can be obtained as efficient as in the traditional face-to-face evaluation [10], allowing to take advantage of this additional benefits without drawbacks.

The use of remote testing on mobile applications such as payment applications is less established [11] than on computers. Therefore, there are fewer guides on how to conduct a remote test on a mobile prototype, specially synchronously [12].

Operating with a mobile phone supposes additional problems compared to a personal computer. Some of this problems are the disparity of mobile operating systems, the lower computing capacity or the greater difficulty to operate simultaneously with different applications.

Many software development companies are developing and improving payment applications following this trend. This work was born as a result of a project in one of them.

Vipera Iberica SL is a young Spanish software development company that focus on working with banks. It is currently developing several payment applications, among other similar projects.

One of these projects is the development of a dedicated mobile payment application for a joint project of Spanish banks.

That is why, when seeking to validate the prototype of this application, the need arises of finding a procedure to conduct remote tests effectively. After discovering that the current literature is scarce on this matter, the idea of this work also arose: To propose a remote usability testing procedure.

1.1 Research Questions

The project should bring solution to the problem of the remote usability testing with mobiles paradigm. Research questions must be established in order of guide and focus the approach of the project. They allow to evaluate also the grade of success of the results.

1. How to perform remote usability testing for mobile payment applications?
2. What tools can be used to perform remote usability testing for mobile payment applications?
3. How can the remote usability testing procedure be adapted to different needs?

The first research question can be answered with the proposal of a remote usability testing procedure. This procedure must meet a series of requirements in order to

answer this question. These requirements are elaborated in detail in the following subsection.

The second research question entails an analysis of the available software and how can it help to the procedure. It is necessary to analyze the usefulness of these tools even though the formulation of the procedure involves the use of specific software.

The third research question can be answered from the analysis generated by the previous two. Although the aim is to propose a valid procedure, its usefulness would increase if flexible tools are offered. These can allow the procedure to be adapted to different case studies.

1.1.1 Procedure Requirements

Due to the breadth of the first research question and the process it is necessary to stipulate certain requirements. In order for the procedure to be considered valid and useful it must meet these requirements:

- The procedure must be simple and accessible for different types of people.
- The procedure must be as non-intrusive as possible.
- The procedure must guarantee that a usability test can be performed correctly.
- The procedure must be easily adaptable to apply in different kinds of mobile application design projects.

The first two requirements aims to get the number of potential testable people to be as large as possible.

A good design should guarantee its use by the maximum possible number of users. It is imperative to try not to exclude anyone when designing regardless of their possible limitations.

Accessibility is essential for developers and organizations that want to create high quality websites and web tools, and not exclude people from using their products and services. Accessibility also benefits people without disabilities [13].

The non-intrusiveness is an especially critical aspect due to the personal nature of the mobiles. If users feel that their privacy has been violated, they may abandon the usability test. This will result in wasted time and also will not guarantee the usefulness of the procedure.

Even if the prototype to be evaluated is oriented to a more limiting target, the procedure must avoid leaving out any potential users of the application. This applies both in terms of personal capabilities and access to technology.

The third requirement seeks to guarantee the effectiveness of the evaluation process. The whole procedure is meaningless if the final objective, which is to perform a usability test, is not met. Therefore, the procedure must allow obtaining valid and useful data during the usability test.

The last requirement seeks to guarantee the generalizability of the procedure. During the project, the procedure will use the same application prototype, but the goal is to be able to use it in other prototypes and projects.

In turn, an effective usability test must be carried out with different users, so it is necessary that the procedure can be easily replicated for the different necessary tests.

1.2 The Project

Based on the current literature and current available tools, a first remote usability testing procedure is proposed that seeks to meet all the objectives and requirements. This procedure will use the prototype of a payments application to perform the usability test.

An experiment consisting of carrying out said procedure with a sample of users of different ranges of age, gender and educational and economic level is proposed. This experiment aims to prove the validity of this first procedure.

Because the prototype to be evaluated is designed to operate in Spain, all the people interviewed will be Spanish. Seeking to guarantee the representation of all national areas, interviewees are selected from different parts of the country.

Interviews for the experiment will include the proposed procedure itself, as well as a discussion about the user's experience when participating in it. With them, the pain points of the procedure are located, as well as the general impressions.

After the first iteration, a second remote usability testing procedure is proposed. This new procedure aims to correct the most conflictive aspects of the first procedure. A new experiment with users is done to test the changes.

The sample of users of this second iteration can be smaller if results obtained during the first experiment are positive. This second experiment seeks to check if the pain points have been solved or mitigated. Additionally, it seeks to verify that these changes do not have a negative impact on the collection and value of the results gathered with respect to the first iteration.

2 State of the Art

This section presents the concepts and technologies necessary to understand and conduct the study.

To provide some context about the topic, a study on the current situation of payment and banking applications in order it's carried on. An analysis of the current trends in terms of payment methods with these applications is also done.

An in-depth study on usability tests is also carried out. Different methods and practices are analyzed are presented to understand the possible benefits and complications that can provide to the remote usability testing procedure.

2.1 Payments Applications

As stated in the introduction, payment applications are a technology on the rise. To understand them correctly, it is necessary to understand their objectives and purposes.

Payment applications allows to make mobile payments.

Mobile payments are defined as the use of a mobile device to conduct a payment transaction in which money or funds are transferred from payer to receiver via an intermediary or directly without an intermediary [14]. Mobile payments are defined also as the act of making payments using mobile devices including wireless handsets, personal digital assistants (PDA), radio frequency devices, and NFC based devices [15]. Mobile payment includes the initiation, authorization, and completion of a single payment [16].

Payment applications can make in remote or proximity payments [17]. Payment applications are usually focused on one of this kind of payments. Depending on the tools and objectives of the payment, they can belong to one or the other. Figure 1 shows the differences between these types of scenarios.

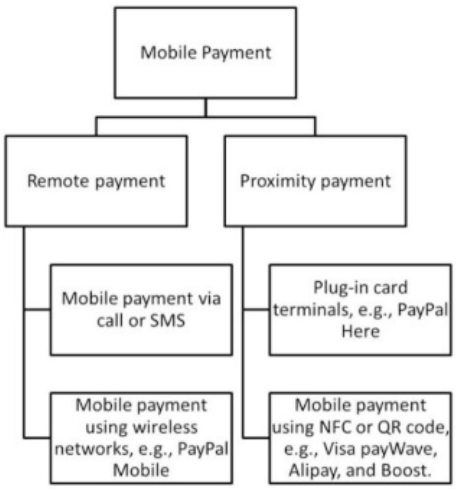

Figure 1: Mobile payment classification framework [18].

Remote payments can be made from virtually anywhere. That is why they need remote means of communication such as the internet or the telephone line.

On the contrary, proximity payments must be made in the same place. They are therefore specially designed for payments in stores. This opens the possibility to additional forms of communication such as NFC or QR. Both technologies will be explained below.

The study of this work focus on proximity payments. This is because proximity payments includes payments in stores and the payment application in develop has this purpose.

Another similar division can be made in terms of some financial aspects: Mobile Money and Mobile Banking [19, 20]. This criteria is compared in Table 1.

Parameter of Comparison	**Mobile money**	**Mobile banking**
Bank account	Does not need bank account.	Does require bank account.
Provider	Financial service provider company.	Bank.
Transactions	Transaction of P2P payments.	All kinds of transactions.
Services	Acts as a Mobile wallet	Acts as Net banking
Banking	Agent banking.	Branch banking of the bank.

Table 1: Comparison between Mobile Money and Mobile Banking [20]

Mobile money applications allows to make payments and operations without having to depend solely on a bank. Because of this, they can allow to operate on different accounts simultaneously even between different banks.

Some of these applications work with their own balance and function as digital wallets. Others allow to link bank accounts and select which one to use when making a payment.

Mobile banking is a service or product offered by financial institutions that makes use of portable technologies [1]. These applications seek to offer a remote and portable service for all banking operations.

These types of applications are not within the scope of this project. This is because the prototype application seeks to be its own independent payment application.

2.1.1 About Payments Methods

As can be seen in Figure 1, the current technologies most used for making mobile payments are the use of SMS, QR or NFC.

These technologies allows to easily synchronize the different devices involved in a transaction. When the devices have the necessary information, they will communicate with the banking and financial entities or systems. This entities will carry out the transaction after receiving the request. When this transaction is concluded, they will again inform the devices involved of the result of said operation.

All these communication technologies allow to share information necessary for processing payments. Alternatively, they can include a link to a server where this information is hosted as a measure of security.

In order to understand how communication is established between devices, it is necessary to understand these technologies in mobile payments.

SMS: Short Message Service, or SMS, is the basic system for remote mobile payments [21]. Mobile payments with SMS requires a communication protocol that allows the exchange of short text messages between two mobile devices [22]. Due to the simplicity of this technology, it has become very popular. Due to being more known and practiced, it is also the method that is often preferred for making mobile payments [21]. However, due to its remote and lightweight nature, this system can pose serious privacy hazards [23].

Figure 2: Example of a QR code

SMS is limited to 160 characters although currently up to 1600 characters can be send in linked messages. This can be an additional problem when transmitting information. If a lot of information needs to be shared, the character limit may not be sufficient.

The arrival of SMS is not completely guaranteed. Between 1% and 5% of SMSs are completely lost [24]. This has led to questioning their use in emergency situations [25].

QR: Quick Response code, or QR codes, work in a similar way to traditional barcodes. QR codes uses a dot matrix to storage information. This information, printed or displayed in a screen can be read and translated with the camera of another device [26]. As an additional advantage over bar codes, is that the code is readable regardless of its position.

QR codes are well suited for mobile proximity payments. Its use is therefore ideal for payments in stores.

Their use has become very widespread, and goes far beyond being limited to mobile payments. The main uses are packaging, mail, magazines and newspapers [21].

NFC: Near Field Communication, or NFC, communicates within a short range to enable data exchange between devices at a distance of a few centimeters [27]. The NFC technology is commonly embedded in mobile phones. Its use for payment transactions is widely used in today's mobile payment systems based on contactless infrastructure [28].

This payment system is considered for mobile proximity payments. Its use is therefore ideal for payments in stores.

NFC has proven advantages [29]. One of them is that its use only requires the installation of a dedicated chip. Additionally, it requires a conscious task of bringing the device closer, which also avoids possible unwanted actions.

Despite this, the system may still be vulnerable. It is possible to eavesdrop from a distance of up to 10 meters [30]. Despite the security of this system, it is necessary to point out its vulnerability to specialized attackers.

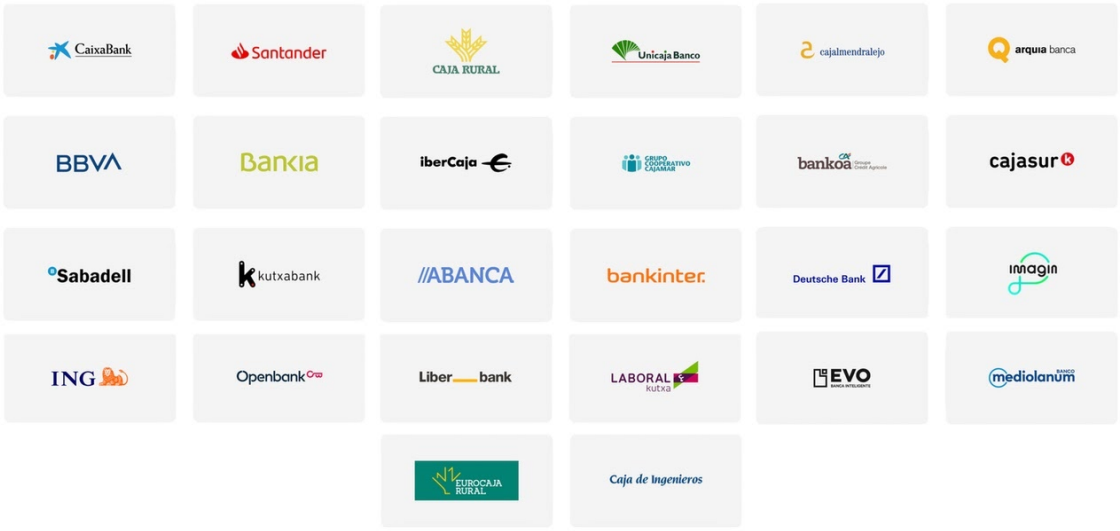

Figure 3: Banks that operate with Bizum [32]

2.1.2 Bizum

In 2016 Bizum was born, as part of Sociedad de Procedimientos de Pago SL. This project was born in response by Spanish banks to the growing boom in new companies and start-ups in fintech.

Bizum operates as a collaborative system for 27 Spanish banks. Participating banks account hold 95% of the Spanish market share [31]. These banks are shown in Figure 3.

Byzum allows monetary transactions between collaborating banks. Transactions are processed from the different banking applications. To carry out a transaction, it is only necessary to indicate the receiver's phone number and the amount of money to be sent.

This makes it a very simple and flexible payment system for the users. This is especially true for remote mobile payments. However, its use is also being consolidated to make mobile payments in stores. This is why the option of a payment application in stores was born.

Bizum is not a proper mobile application, but rather an extension of the banking applications of the banks that make up the project.

Since its creation, its use is in full expansion. In 2019, 6 million users were registered [33]. By 2020 this number had grown to 10 million [34]. So far in 2021 the number has grown again to 15 million [35].

2.2 Usability Testing

According to ISO 9241-210 human-centered evaluation provides an effective means of minimizing the risk of a system not meeting user or organizational needs. It also take place as part of the final acceptance of the product to confirm that requirements have been met [36]. Along with usability evaluation, it aims to reveal usability problems,

to find ways to improve the product and, thereby, to help the development team to fulfil the user requirements [37].

Usability testing, as part of the usability evaluation process, is the activity that focuses on observing users working with a product, performing tasks that are real and meaningful to them [38]. The primary goal of usability testing is to improve the usability of the evaluated product. For this, its needed to analyse the data, diagnose the usability problems, and recommend changes to fix the problems [39].

The most widespread way to perform usability tests is in a laboratory [10]. Starting from a controlled scenario, moderators can give users the devices they need for the assessment. By having its own facilities, it is often easy to observe user performance from a nearby room. In addition, it is easy to have all the engraving and data collection tools that are required.

However, as already explained in the introduction, it is possible to perform usability tests without being in the same space with the user. This is what we mean by Remote Usability Testing.

While the goals and benefits of usability testing are clear, a study of different types of usability testing may be necessary to understand the study problem.

In addition, knowing different techniques available is necessary to be able to formulate a remote usability testing procedure that is adapted to the objectives of the project.

2.2.1 Think-Aloud Protocol

One of the most used techniques during usability test is the think-aloud protocol [39]. During a think-aloud study, participants talk about what they are thinking as they interact with the system [40]. This technique not only allows to better understand the flow of thoughts of the user but also allows to identify patterns in the interaction that can go unnoticed only by analyzing the interaction itself.

There are different ways to put this technique into practice. The two most popular today are concurrent or retrospective [40].

Concurrent thinking-aloud (CTA) is the most popular approach due to its fastness, efficiency and ease for users to relate to [37]. With this technique, users vocalize their thoughts during the execution of tasks. This process can slow a little bit down the interaction a bit as users often formulate complete sentences before continuing with new actions.

With Retrospective thinking-aloud (RTA), they instead perform the task without verbalizing their thoughts. It is after the interaction that users are asked to share their thoughts. To do this, a recording of their interaction is usually shown [40]. This technique is usually longer due to the need to review the interaction. While the user is guaranteed to be fully focused during tasks, some context may be lost when remembering thoughts.

Despite these differences, multiple studies have proven that the effectiveness of both techniques is practically similar. Users subjected to CTA tend to have a higher number of errors. Still, this does not affect usability flaw detection. On the contrary, users who perform RTA usually feel more disturbed than performing CTA [37].

Although both methods are valid, the project will focus on the use of CTA. This will greatly facilitate the procedure by not needing to review the interaction. In addition, since the greatest simplicity for the user is sought, this technique can be less confusing.

2.2.2 Remote Testing

This technique has certain advantages and disadvantages compared to studying in a laboratory according to previous studies [7]. It is important to be aware of these points. With them, it will be possible to exploit the benefits that this technique can offer.

A clear advantage of this technique is that a laboratory is not required. This entails a reduction of costs.

It also saves time and travel. This includes both the users and the research team. This time reduction brings with it many other advantages. With less time needed for each usability test, it is possible to perform more of them. This makes it possible to increase the number of users to be tested. Another alternative is to perform longer and more in-depth usability tests to detect more usability problems.

Finding people to test it is easier also because they will not be deterred by having to travel. Not relying on the user to travel to the laboratory also makes it easier to evaluate users in different locations. It also allows the research team not to be dependent on being in a specific location.

Allowing the user to operate from a friendlier environment can also have benefits on the quality of the data obtained. If the user is comfortable, it is possible that their responses will be more natural. Additionally, it will be more difficult for the user to become stressed or saturated during the test.

Finally, this technique allows to automate some processes. On the one hand, it is possible to automate the evaluation of the interaction results by allowing the user to perform the tasks alone. It is also possible to automate the search for users via the Internet. This again can lead to a reduction in time and resources.

On the other hand, knowing the disadvantages can help to resolve flaws in the procedure.

The main one is the increased difficulty in collecting user data. With no direct observation, some problems may remain undiscovered because gestures and facial expressions of the user are missing.It's harder to know also whether the test persons really test the app carefully and seriously.

In general, it is more difficult to interact, interview, train and observe the participants. This also means being even more careful when working with and viewing confidential information.

Another major difficulty is on the technical side. The prototypes must be adjusted to be able to be used remotely. This may require some logistics in order to distribute hardware or offer support for installing software.

In addition, technical problems are more likely to occur throughout the process. Support to solve these problems can be limited and complicated.

2.2.3 Remote usability testing variances

The remote usability test can in turn be divided into two different categories: Synchronous and Asynchronous [41].

A synchronous test is one in which the user is in contact with the moderator and the team by calling. Similar to the laboratory study, the team can obtain the data in real time.

On the other hand, an asynchronous test is one in which the user performs the tasks without being on call or the moderator being present. Data is usually collected directly from user interaction.

The asynchronous system can facilitate the study in terms of time. This in turn can allow many more users to be evaluated. For the nature of this project, the information obtained synchronously is considered to be of greater value.

It can be expected less information to be collected asynchronously. This is due to the static nature of the test. Being able to redirect the conversation or the questions through an moderator can add a lot of value. Despite this, no value comparisons have been found regarding the usefulness of these data with respect to a laboratory test.

While the asynchronous system can be of great value for specific cases, the synchronous process can be the solution for the majority of case studies. For this reason, the synchronous remote testing is the one that will be proceduralized.

It should be remembered that the information that can be obtained through a synchronous remote usability test can be of the same value as that obtained in a laboratory study [10].

2.3 Evaluation Techniques and Questionnaires

To evaluate the usability and the validity of the procedure, it is advisable to use metrics of proven effectiveness. Using simple questionnaires, it is easy to obtain data that allow the extraction of qualitative or quantitative information.

Quantitative information is of special interest when it can be compared. This information allows to make statistical analysis of different iterations of experiments.

2.3.1 SUS Questionnaire

> *"Usability is not a quality that exists in any real or absolute sense. Perhaps it can be best summed up as being a general quality of the appropriateness to a purpose of any particular artefact".*
> John Brooke, SUS: A quick and dirty usability scale[42]

This is why usability must be measured based on certain criteria. The System Usability Scale (SUS) questionnaire has established itself as one of the most widespread tools to measure "usability" [43].

SUS consists of a series of 10 statements. These cover different subjective aspects about the system to be analyzed. Each of them is evaluated numerically according

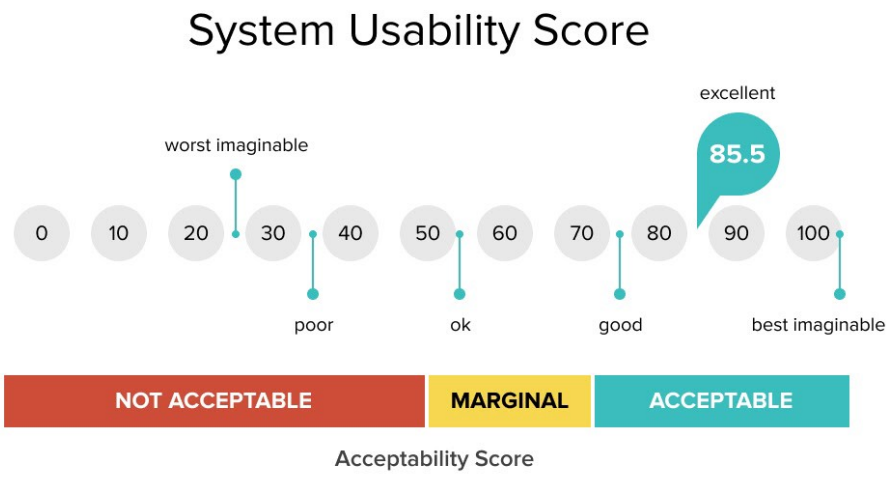

Figure 4: SUS score Acceptability Scale [44]

to the degree of agreement of the user with respect to the statement.

Affirmations alternate positive and negative statements as a measure of ensuring the user's awareness. Using a simple algorithm, it is possible to obtain a final value ranging from 0 to 100. The higher this value, the higher the user satisfaction.

This final value does not provide definitive information. Instead, it must be subjected to analysis to understand its value. In any case, there is some agreement as to whether the score can be considered positive or not. Figure 4 shows an interpretation of the acceptability of this result.

The validity of the SUS questionnaire as a tool for the study of usability is well proven. However, it should be noted that the formulation of this is in English.

Because this study takes place in Spain, it is important that this tool is understandable by users. The paper "Spanish Version of the System Usability Scale for the Assessment of Electronic Tools: Development and Validation [43]" offers a proved Spanish translation.

2.3.2 NASA-TLX Questionnaire

To evaluate the procedure we want to evaluate the satisfaction and comfort of the users when being led through the process. Therefore, what needs to be evaluated is the procedure and not a product or service.

Effort can be evaluated in different ways and can be originated by different aspects. Nasa-TLX (Task Load Index) questionnaire allows to evaluate the workload in a process. It also measure the different aspects that contribute to the general effort.

But first it is necessary to be clear about what is meant by workload. It is the level of deliberate and conscious control of information necessary for behavior to occur. Workload can be described also as the portion of the operator's limited capacity required to perform a given task [45].

Workload is divided into different aspects. Each of them evaluates a component

that can negatively affect the user. The defined aspects are the following [46]:

- **Effort:**. Mental and physical effort that the subject has to make to obtain her level of performance.
- **Mental demand:** Mental and perceptual activity required by the task. This can include thinking, deciding, calculating, remembering, looking or searching.
- **Physical demand:** Physical activity required by the task. This can include pressing, pushing or turning.
- **Temporary demand:** Temporal pressure level felt. Ratio between the time required and the time available.
- **Performance:** To what extent the individual is satisfied with his level of performance.
- **Frustration level:** To what extent does the subject feel insecure, stressed, irritated or unhappy during the performance of the task.

The formulation of this questionnaire and its interpretation are a bit complex. In the first place, the user is presented with all the possible permutations of each of the aforementioned aspects faced with another.

The user must state for each pair of aspects which is the one that has led to the greatest workload. The objective of this part is to assess which are the most controversial and weighty aspects for the total load as a whole.

The user is then asked to weigh how much each aspect has contributed to the total load. To do this, the user must assign a numeric value. The greater the impact of this load, the higher the result.

Ideally, the weighting of aspects is done without numerical values, to try not to mathematize the process for the user. For this same reason, the weighting is carried out after comparing aspects. If the reverse is done, the user may simply compare numerical values and respond accordingly.

Once all the data has been collected, the scoring algorithm is responsible for obtaining a load value for each aspect analyzed. This value will take into account how many times the aspect that most influences the load has been considered in each comparison. Also, the value will be adjusted based on the results of the weighting for that aspect.

In this way, it is possible to analyze the severity of each aspect in the total workload. A final score for the complete process is also obtained based on the individual values of the evaluated aspects.

Similar to the SUS questionnaire, the numeric values obtained through this questionnaire do not give absolute information by themselves; instead, a study needs to be carried out to analyze the meaning of these values.

It is more complicated to establish acceptability ranges due to the wide variety of tasks measured with this questionnaire. This is also complicated because it admits a much wider range of scores than SUS.

Percentile	Score
Min	6.21
10%	26.08
20%	33.00
25%	36.77
30%	39.45
40%	45.00
50%	49.93
60%	53.97
70%	58.00
75%	60.00
80%	62.00
90%	68.00
Max	88.50

Table 2: Percentiles of Global NASA-TLX Analysis [47]

It is possible to establish an assessment by comparing with results obtained in similar tasks. According to a study by Rebecca A. Grier [47], scores ranges from 6.21 and 88.50. Tasks labeled as "Cognitive Tasks" ranges from 13.08 to 64.90. This study gives also percentiles scores that allow to rank the results in comparison with others. Table 2 shows the global percentiles comparison for NASA-TLX scores.

Numeric values can be also useful for comparative studies such as the SUS ones. This is of value in analyzing the differences between different iterations of an experiment.

With this information, it is possible to detect components that are having a negative impact on the overall experience of the process. By detecting these points, it is possible to evaluate how to mitigate them.

3 Methods

The evaluation tools are set in this section. These will allow us to analyze the degree of success in achieving the set goals of the study. To select these evaluation tools, it is necessary to take a look at the research questions to be answered.

How to perform remote usability testing for mobile payment applications?: This answer of this question is directly linked to the validation of the procedure. The procedure is proposed to provide an answer to this question. This question is answerable if standardization of a remote usability testing procedure is possible.

It is necessary to stipulate success criteria in order to validate the proposal of the procedure and to be able to consider this research question answered. A number of requirements were stipulated in the introduction. Four properties are extracted from the requirements.

Meeting these 4 properties will make it possible to evaluate whether the procedure is valid and to answer this question. The establishment of these properties and how to evaluate them is defined in this section.

What tools can be used to perform remote usability testing for mobile payment applications?: A preliminary study of the available tools must be carried out before proposing the remote usability testing procedure.

This analysis will be conducted in the next section while setting up the experiment. The experimentation in turn allow to evaluate the suitability of the tools chosen and analyzed. The results obtained after the experimentation allow to extend the analysis to answer this question.

How can the remote usability testing procedure be adapted to different needs?: This question is solved in a similar way to the previous one. The study prior to the experimentation allows to know the available tools and which of them are possible alternatives to the proposed procedure.

In the same way, the experimentation allows to obtain information to be able to elaborate a discussion on the possibilities of adaptation of the remote usability testing procedure.

3.1 Evaluation of the Procedure Requirements

To properly evaluate the effectiveness of the procedure, and answer the first research question, it is necessary to establish an evaluation system. For this evaluation it is necessary to take into account the requirements that the procedure is expected to fulfill.

Starting from these requirements, four properties to be evaluated can be extracted:

- Simplicity
- Intrusiveness
- Effectiveness
- Viability

The remote usability testing procedure may be considered valid if it has these properties.

Since two iterations of the experiment are to be performed, it is necessary to analyze the individual and comparative results.

Some basic objectives for the validity of the procedure can be defined. Even so, the second iteration is expected to improve the values obtained even if the initial proposal turns out to be valid.

In the event that the second procedure turns out to be invalid, it would be necessary to develop new iterations to achieve this goal. It would also be necessary to continue the study if the second procedure turns out to worsen the values obtained with the first procedure.

3.1.1 Simplicity

The procedure must be simple and accessible for different types of users.

This aspect assesses the ability of users to complete the procedure. It must be guaranteed that any user is able to participate in the procedure. Developing a procedure that does not meet this objective may mean that certain user profiles cannot be taken into account.

In addition to this limitation, performing the procedure to be unable to complete it entails a waste of time. This loss can also have major consequences such as economic or trust.

The heuristics defined by Jakob Nielsen[48] are widely known in the field of interface design [49]. It is not easy to make a direct parallelism with these heuristics because what is being designed is a process. However, it is necessary to keep these properties in mind in order to guarantee the simplicity of the procedure.

The way set to evaluate this is simple and is to get as many users as possible to be able to complete the entire procedure. As a goal to be achieved, the effectiveness of the interviewees is expected to be at least 95%. This means that interviewees are able to complete the entire procedure without becoming blocked. Accessibility must also be taken into account to ensure this goal.

In addition to complying with this percentage, the opinion of interviewees in this regard must be taken into account. Although all of them manage to finish the procedure, if the feedback received shows a large number of difficulties, the remote usability testing procedure could not be considered valid either.

In the final discussion, it is needed to ask interviewees their general opinion and their biggest problems in order to assess whether the procedure is simple enough.

In this final discussion the Nasa-TLX form is also presented to interviewees. This can allow to detect components of the workload that are too high.

If any of the criteria of the Nasa-TLX questionnaire reaches very high values, it would be necessary to improve that aspect to give the procedure as valid in terms of simplicity.

The use of this questionnaire also offers the possibility of making a simple comparison between different iterations of the procedure. If the values obtained in

any of the aspects worsen in a revised version of the procedure, this would rule out its improvement and its validity.

Therefore, to ensure that the process is simple, a 95% effectiveness rate must be achieved by the interviewees in carrying out the process. In addition, the feedback received by the interviewees should indicate that there are no major difficulties.

3.1.2 Intrusiveness

The procedure must be as less intrusive as possible for the user.

To ensure user comfort and confidence, the intrusiveness of the procedure must be minimized.

The need to download external applications or files can be considered intrusive. This not only makes the procedure more complex but can also generate mistrust by not knowing the origin or impact of these downloads. If done during the call, this can be worse, as the user may feel rushed, which can lead to frustration.

It can also be intrusive to request personal information, especially regarding accounts, emails or phone numbers. The less information of this type is required from the user, avoiding the possibility of making him want to not continue with the procedure.

It is possible that in the search for usability test data, information that can be considered private is also required. This makes it more important not to help scare the user away.

Compliance with this objective can be assessed if interviewees do not decide to abandon the procedure because they do not want to share necessary information. However, this criterion may not be sufficient due to the multi-layer nature of the study. Users may feel more inclined to share this information because of the special nature of this study than in a regular study case.

This is why this criterion should be evaluated mainly by asking interviewees for their opinion. If the feedback received insists on the discomfort of needing to share certain information, an alternative, if any, should be sought.

3.1.3 Effectiveness

The procedure must guarantee the correct performance of a usability test.

This criterion seeks to validate the usefulness of the procedure. As much as the other criteria are met, the procedure is meaningless if it does not fulfill its basic function: Detect usability failures.

It is in this aspect that the prototype to be evaluated takes on special importance. In the first place, to guarantee the effectiveness of the procedure, it must be possible to obtain data and useful information for the improvement of the prototype.

This aspect is also of special comparative importance for the second iteration. Once the changes in the procedure have been stipulated, it must be evaluated that they do not have a negative impact on the quality of the information obtained.

For this comparative study, it is not only possible to compare the problems reported by users but also the results of the SUS questionnaire. If the results of

this questionnaire vary widely in different iterations, they could indicate that the procedure is impacting and biasing the collection of information.

If this happens, a detailed study would have to be made, since a direct relationship cannot be made as to whether better results in the SUS questionnaire imply a better or worse experience in the procedure and vice versa.

It is important to note that, once the objective of obtaining useful information through the usability test has been fulfilled, the objective of the changes on the procedure it won't be to improve data collection. Nor does it seek to obtain information more effectively than in a laboratory study.

The objective to achieve from there is to try to improve and smooth the user experience during the procedure. Even so, it is possible that this improvement of the experience could lead to an improvement in the results.

3.1.4 Viability

The procedure must be easily replicable to apply in different mobile application design projects.

Finally, to guarantee the usefulness of the procedure, it must be able to be easily replicated. This includes both by the team and by the users. Again, it is important not to leave user profiles out of the possible study.

The very approach of the experiment should take this point as valid. Repeatability by users is evaluated at the point of effectiveness.

On the other hand, the replication by the moderator is evaluated with the completion of the entire iteration of the experiment.

4 Construction and Design of Experimentation

This section presents all the tools that will be used to conduct the experimentation. These tools of different types are the ones that allow the remote usability testing procedure to be carried out.

It is necessary to specify the scope of the work at a multi-layer level, to understand better how to conduct the experiment.

In order to know the available software tools, it is necessary to carry out a prior analysis. This allows to know the benefits of each one of them. Knowing this also allows to choose those suitable for the remote usability testing procedure and helps to answer the research questions.

It is also necessary to know the prototype on which to perform the remote usability test. This is important as it determines the information to be obtained during the test. The tasks of the usability test need to be explained in this section too.

Finally, to establish the criteria for the selection of users is also needed. In order to obtain valid results from both the application study and the procedure, the sample of users must be chosen with certain criteria. This includes the number and characteristics of the users.

4.1 Multi-Layer Study

An added complication to this project is its multi-layer aspect. The objective of the project is to propose a remote usability testing procedure, but this entails a two-level study. On the one hand, the proposal, study and validation of said testing procedure is carried out. On the other hand, to carry out this procedure, a usability test itself must also be carried out on the payments application, seeking to find flaws in it's design.

This two levels can lead to confusion both to understand the work itself. It can be also confusing for the users themselves during the experiment. So, due to the multi-layer nature of the project, it is necessary to clarify some concepts to be used and avoid confusions.

- **Procedure:** The remote usability testing procedure (or just procedure) includes the entire proposed method for conducting the remote usability test. The procedure its improved during the project experimentation in two iterations of the experiment.

  Each of this iterations begin with a procedure proposal and tries to prove the validity of the remote usability testing according with the requirements established. The different iterations of the procedure during the analysis are called "first procedure" and "second procedure".

  The "final procedure" is the solution proposed to solve the first research question as a conclusion of this project.

  The remote usability testing procedure consists of the invitation of the user, the call setup and introduction to the project, the setting up of the mobile and

Figure 5: Each interview includes the full remote usability testing procedure (phases 1 to 5) and the evaluation of the procedure itself (phase 6)

the remote usability test itself. This usability test includes the execution of the proposed tasks, the evaluation of the experience and a short discussion about the experience with the prototype. Finally the user is dismissed and the procedure ends.

The different phases of the procedure are explained in full detail in the Solution and Results.

- **Interviews:** The different sessions during the experimentation in which the entire remote usability testing procedure is carried out. The interview includes also a evaluation with the user about the procedure itself. This includes the evaluation questionnaire for the procedure (NASA-TLX) and a conversation about the experience of the user during the course of the procedure.

- **Experiment:** Two iterations of experimentation are carried out to evaluate the different versions of the remote usability testing procedure. The entire process of interviews testing one procedure proposal is what is called as experiment.

  The first iteration of experimentation (or first experiment) seeks to find pain points in the first proposed procedure. The second experiment seeks to test the validity of the changes in order to propose a final remote usability testing.

  The experiments and their results are explained in full detail in the next section.

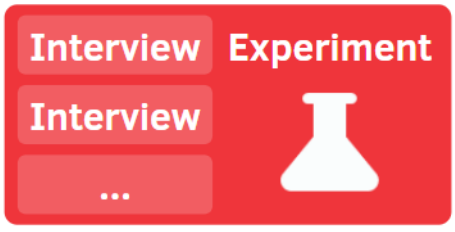

Figure 6: Each experiment correspond with the full iteration of interviews evaluating the same procedure

4.2 Software and Tools

The remote usability testing procedure must provide a solution to different needs. Each of them can be solved with software solutions. Therefore, it is necessary to carry out an analysis of each of these needs, as well as the possible tools with which to solve them.

Software tools should solve each of the needs but they must be simple and accessible too. The chosen solutions should also not be very intrusive or entail costs. They should avoid unnecessary records and they should be executable on most hardware available to the users.

This section is divided into different sections to facilitate its organization: Call, Prototyping and Questionnaires.

Each subsection discusses the tools available to address different remote usability testing procedure needs.

4.2.1 Call

While on phone call or video call, the moderator can conduct the usability test remotely. Although the procedure begins with an invitation and the presentation of an Informed Consent to sign, the establishment of the call is surely the most critical aspect. To perform the usability test synchronously, it is necessary to share a call with the interviewed user.

The suitability of the calling tool is determined based on the following aspects:

Video-call: Ideally, this system should allow the availability of a camera. In this way, it is possible to see and analyze the gestures of the user in order to better understand their emotions. Also, allowing the user to see the moderator will facilitate communication and can help make them more comfortable.

Recording: The ability to record the conversation can be helpful. This can make easier to analyze the data again and again in the future. Also, being a small team can allow the team to take the worry out of taking notes during the interview and do it thoroughly later.

Call recording can be done with additional software if the call tool does not allow it, but unifying tools is a positive aspect in order to facilitate the procedure.

Installation: In order to be accessible to users, avoiding having to download and install the application can facilitate the procedure. This is possible using functionalities already integrated into the mobile. This is also possible if the tool can be run from a browser.

On the one hand, people with less technological prowess or more insecure avoid having to deal with the application download process.

On the other hand, having to download applications, especially if they are heavy and consume many resources, can leave potential users out of the study as they do not have a device that can run the tool.

Registration: Similarly, it can be beneficial to avoid tools with complex registration systems. Less skilled users may get confused or uncomfortable during this

registration. These users may eventually decide to abandon the procedure. The abandonment rate can be as high as 11% if users have to register for an online service or they are asked for too much information [50].

Therefore, the easier it is to register, the better. If it is possible to use the system without any registration, even better.

Screen share: A very positive aspect may be the ability to share a device's screen. With this option, one of the devices in the call can give access so that the invitees to the call can see the screen of the user who is sharing. This can be useful for sharing documents, instructions, or even showing the prototype.

Another positive aspect is being able to join the call from the mobile device, share a screen and run the prototype. This also allows the moderator and the team to follow the user's interaction. Being able to see this interaction live also makes it easier to help the user in case of encountering problems. These problems are not limited only to the performance of the usability test user tasks. Screen sharing from the mobile can allow the moderator to support the user to perform tasks related to the procedure beyond the usability test.

Additionally, the ability to give control over the shared screen can be very useful. In this way, a member of the call can use their mouse, keyboard or mobile controls to perform actions on the shared screen.

This option can allow the moderator, for example, to configure the interviewee's devices. Another interesting option would be to run the prototype from an own device and give the interviewee access to operate on it, avoiding him having to download or run the prototype.

Interference with the usability test: Since the procedure focuses on mobile devices, it is important to ensure that the call setup does not affect the use of the prototype.

To guarantee the interaction naturalness, the user tasks presented to the user must be carried out on a mobile device. In turn, this interaction should not be impaired by the establishment of the call.

Although it is possible to run the prototype on another device and show the simulation directly on the user's mobile via streaming, this interaction must also be realistic and not affected by external elements.

Device support: As explained in previous points, it may be necessary and convenient to be on call with both mobile devices and computers.

It may be easier for the team to have a computer ready to easily carry out the various tasks in the procedure.

Interviewees may have to join the call from their mobile. This way the user can run a simulated prototype or share a screen to be able to follow his interaction in the usability test.

In addition, it may be easier for the user to have a separate device to maintain the call. This can be a computer, a tablet or even another mobile device. In this way, the execution of the prototype and the tasks are easier to maintain without intrusions from the call tools.

4.2.1.1 Available tools

Having detailed all the aspects to be evaluated, it is possible to carry out the analysis of the different tools available for the establishment of the call.

Figure 7: Whatsapp logo

Telephone Call: The telephone call is surely the simplest system for most users to maintain a call. As positive aspects it can be pointed out that any user is used to this system. Additionally, no additional software is necessary beyond the mobile.

Many devices allow the option to record the call, but it is not guaranteed and users do not have to be used to it.

This system does not allow to be seen through a camera. Nor does it facilitate per se other aspects mentioned such as screen sharing or being able to be done from a computer. These aspects should be solved with additional software.

Whatsapp: Online messaging service of Facebook, Inc.

WhatsApp is a messaging service that allows you to have conversations in chat and on video. These conversations can be direct between users or in groups.

This application is widely spread. Although it is linked to a telephone number, it is expected that the vast majority of users are familiar with this system.

WhatsApp offers the option of making video calls very easily. However, it does not allow screen sharing or control.

There is the option to open WhatsApp from a computer both from the browser and from the dedicated application. This allows (from the desktop application) to be able to establish calls simultaneously on mobiles and computer.

Telegram: Online messaging service of Telegram FZ-LLC.

Telegram offers a service very similar to WhatsApp. Its use is less widespread and it is more likely that users do not have this application or have linked their numbers.

In return, Telegram offers a couple of advantages. One of these is the possibility of contacting users through a username, apart from the phone number. In terms of privacy, this can be less intrusive for those who are wary of sharing their phone.

Like WhatsApp, it is possible to open Telegram from the computer in a browser or in its own application. Unlike WhatsApp, Telegram allows screen sharing from a mobile device or tablet or from the desktop application.

Figure 8: Telegram logo

Telegram doesn't have the option to share controls either.

Hangouts: Google Video Calling Tool. This system is linked to the Google account system.

This system allows the establishment of calls in a not very complex way. Calls can be made individually or group rooms can be created.

In the case of the individual call, another users can be called using their Google name, their email or their phone number. The invited user will need to identify themselves with his Google account to join the call. If the user does not have a Google account, registrating and creating a Google account is needed.

Figure 9: Google Hangouts logo

For group calls, it is possible to create a permanent room and join it through a link or an invitation. In both cases, the user must also identify himself with his Google account or register.

In the case of wanting to join the call from a computer, this can be done directly from the browser, avoiding unnecessary downloads beyond the browser itself that all users are assumed to have.

When joining from a mobile or tablet, it is necessary to download the Hangouts application, so in this case the user would necessarily have to download something.

If the user is already on a call from a computer, the moderator could support the user during the call to help him with this download if the userhas problems. This support can be done only with the user's instructions or by asking him to show the mobile through his camera. The latter is not a very efficient system, but it can be used to solve punctual problems.

Once in a call, the system allows screen sharing from a computer but not from a mobile or tablet device. However, sharing of controls is not allowed in none of them. In this way, the moderator can show his screen and show documents or help the explanation of instructions. The moderator could show the prototype but cannot give permission to the user to operate with it from his mobile.

The user also cannot share his screen to show his interaction with the prototype. Therefore, other software would be necessary to be able to follow the user's flow.

The user could neither have access to the controls of the moderator's computer or mobile, so the user should be the one to run the prototype from their mobile.

A final positive aspect of Hangouts is its free policy. The creation of an account, the download of the apps and the calling system carry no cost or limitation.

It should be noted that, by depending on the Google account service, Hangouts offers integration with other potentially useful services such as Gmail, Drive or Google Calendar.

Hangouts is in transition to unify with Google Chat.

Zoom: Similar video calling system of Zoom Video Communications.

Figure 10: Zoom logo

Zoom works quite similar to Hangouts. Despite not having the backing of a huge company like Google, this application has gained quite a bit of popularity due to the exceptional situation of the COVID-19 pandemic.

Zoom can be run on a computer from the browsers. It also has a desktop application but its use on mobile or tablet requires the download of its own application. All of this works like in Hangouts.

An advantage over Google's system is that Zoom does not require user registration or identification. It's necessary to have an account for the creation of a room, but this account can be free and only the host should have it. In this case, it would be the moderator, which is not a problem.

It is possible to join the call from other devices on a computer, tablet or mobile through a link or invitation. Zoom allows screen sharing from the computer like Hangouts, but it also allows screen sharing from the mobile. In addition, it also allows to offer control of the screen, but only from the computer.

This opens up new possibilities for the remote usability testing procedure. One is the option to view the user interaction from their mobile shared screen. On the other hand, it opens the option to run the prototype from the computer, share screen and control so the user can use it from the mobile.

In case of sharing the screen with the mobile if there is a small problem to point out regarding the execution of the prototype. Zoom allows the user to keep the call in the background, so that the mobile can be use outside of the application. To manage screen sharing options, a small menu remains visible in the foreground. This can be superimposed on the prototype at the time of execution and muddy the interaction to some extent.

These controls are also visible on the computer and tablet. In fact, Hangouts also has similar controls on the computer. In any case, when operating on a large screen, their impact on a computer is minimal. However, it is necessary to point out this complication in the mobile case.

Regarding the payment policy, it is more limiting than in Google's solution. Although the system is free, there is a limitation for group calls. A limit of 40 minutes is established for the call in case the call has 3 or more guests (up to a maximum of 100). This limit is maintained even if the call is again only between two people when others users leave it.

This can be a problem if the moderator wants to be on a call with one device and two by the user (their mobile phone and their computer or tablet).

Zoom offers a premium system to avoid this. Through a subscription, this time limit on group calls is eliminated.

Premium license offers other advantages such as calls with more guests, call transcription, cloud storage or the use of a corporate image. It should be noted that all these options can be useful, but especially transcription. Not only is it another accessibility tool, but it can facilitate obtaining interview data by directly transcribing the dialogues into text.

Figure 11: Microsoft Teams logo

Teams: Work team management service of Microsoft.

Unlike Hangounts and Zoom, Teams offers a more comprehensive system. Its objective is to offer a complete communication platform for professional teams. This not only includes calls but an integrated calendar, the creation of groups and rooms, integration with Outlook and others.

Although Google has many of these services built into its system, it does so through different applications. Microsoft Teams on the other hand offers to manage everything from a single application.

In this case, the user must have an account that has been integrated into the Teams work team they want to join. An easy option may be to create a dedicated account for him rather than invite him to an existing account.

Although Microsoft Teams has premium services with several additional options, the free system serves to meet the basic objectives of the procedure. Calls have no limitations and are free.

Like the other options analysed, Teams allows to join calls through the browser on a computer. There is also a more complete dedicated desktop application. However, to be able to do it from a mobile or tablet it is necessary to download the dedicated application.

Teams offers the option of sharing screen both on mobile and tablet as well as on computer. Controls can only shared on the desktop and tablet version. This enables the aforementioned options as Zoom. Teams actually offers the option for multiple users to simultaneously control the shared screen. Although this is a positive aspect, it is not of interest to the proposition of the procedure.

It should also be noted that, as in Zoom, the Screen Sharing menu can obfuscate the iteration a bit from the mobile.

Finally, it should be noted that Zoom has been the tool used to conduct remote tests in 2020 according to the 2020 Design Tools Survey [51].

Slack: Team Communication service of Slack Technologies.

Slack offers a service quite similar to Teams but more oriented to the establishment of chat rooms, being lighter.

In a similar way to the previous applications, it is executable from the browser from a computer, also offering an alternative desktop application. Downloading is required to run on mobile devices and tablets.

Similar to Teams, the invitation to the work team is required to be able to participate in the same calls and chat rooms.

It offers easy integration of external services such as Google Drive or Office 365.

Individual calls are free and unlimited. In order to make group calls, it is necessary to purchase one of the premium packages.

Group calls allow up to 15 people with them. The paid service also guarantees the permanence of written messages, since the free version only keeps the last 10,000.

The option to share screen and control also exists but it is only offered within the premium packages. In addition, this option is limited to computer use and is not allowed on mobile devices.

Although it is lighter to use than Teams, which may be easier to use in the procedure, Teams does not require subscription services to be able to make group calls. In addition, the inability to share screen from the mobile makes it a less attractive option than Zoom or Teams.

Figure 12: Slack logo

Discord: Messaging service of Discord Inc.'s with high focus on Gaming in a kind of social network.

It offers a communication system very similar to Slack. Its structure in channels and groups is very much based on the structure of Slack. In fact, it is quite known for being "Slack for Gamers" [52, 53].

As an advantage over Slack, it allows the option of making direct calls without the need to have been invited to a workgroup. In addition, the invitation to groups is made directly with respect to the registered accounts. In any case, the user needs to be registered in the system, unlike Zoom.

Figure 13: Discord logo

Like all the previous ones, it allows its use from a computer browser. There are also dedicated desktop applications and applications for mobile and tablet.

In the case of Discord, the screen sharing option is free and is available on computer, tablet and mobile. Discord allows multiple users to share screen simultaneously on the same call as an added bonus. The rest of the applications, however, only allow one user to share a screen at a time.

Share controls option is not allowed on Discord.

Discord offers premium packages. For the project proposition they are not of interest in general. However, it should be noted that this service also offers the option of being able to share a screen with a higher resolution and share much larger files.

4.2.1.2 Comparison

The comparison of the basic characteristics of interest of the mentioned applications is shown in Table 3. Limited functionalities are those that, although they are viable for free, need a subscription system to be fully realized. "Blocked" refers to those that necessarily need to purchase a premium package.

On the other hand, in the comparison of computer or mobile sharing, it refers to the ability to share screen and / or controls. In the case of "Both", it refers to being able to share both screen and controls. "Multiple" refers to the possibility of being carried out by several users simultaneously.

	Video Call	Group Call	Computer Sharing	Mobile Sharing	Recording
Phone Call	No	No	No	No	Yes
Whatsapp	Yes	Yes	No	No	No
Telegram	Yes	Yes	Screen Sharing	Screen Sharing	No
Hangouts	Yes	Yes	Both	Screen Sharing	Yes
Zoom	Yes	Limited	Both	Screen Sharing	Yes
Teams	Yes	Yes	Both	Both. Multiple.	Yes
Slack	Yes	Blocked	Blocked. Both.	Screen Sharing	No
Discord	Yes	Yes	Multiple Screen Sharing	Screen Sharing	No

Table 3: Functionalities comparison of the applications studied for call establishment.

	Registration	Invitation	Installation
Phone Call	Not needed	Not needed	Not needed
Whatsapp	Needed	Not needed	Needed
Telegram	Needed	Not needed	Needed
Hangouts	Needed	Not needed	Needed
Zoom	Not needed	Not needed	Needed
Teams	Needed	Needed	Needed
Slack	Needed	Needed	Needed
Discord	Needed	Not needed	Needed

Table 4: User requirements for the applications studied for call establishment.

Table 4 shows the access complications for the user. It compares if it is necessary to register an account to access the system. Keep in mind that even so, there are services such as WhatsApp or Google that are widely extended. The popularity of the service may imply that more or fewer users already have accounts on those services.

The Invitation column compares whether it is necessary to invite the user to a dedicated work environment of their own. On the contrary, if it is not necessary, it is possible to establish a call with the user simply by accessing him or allowing him to join the workgroup.

Finally, it is compared whether it is necessary for the user to download and install their own application for use on mobile phones. In the case of operating on a computer, all services allow operating from the browser even if they have desktop applications. Direct telephone call is the only one that does not allow direct operation from the computer.

4.2.2 Prototype: Design and Execution

In order to carry out the usability test, the user must be able to use a prototype of the application. As previously mentioned, the goal of the usability test is to detect design flaws early in the development process. This is why it is expected to use a prototype.

It is still possible to use a functional application, specially in advanced phases of the development process. In any case, operating on an own application can make how to run vary greatly. This is why, although this possibility will be included in the discussion, the procedure will focus on the use of a prototyping tool.

Currently there is a wide variety of prototyping software. These tools offer a framework for the design of prototypes. In addition, most of them not only offer tools for design, but also the simulation of the designed prototype. In this way, simple operations can be defined to allow navigation between screens.

There is also a certain facility to be able to import projects between some of these software. However, for the proposition of the procedure, the design framework is not the most important thing. The most relevant aspect for this study will be the ease of these tools to allow sharing and running the prototype on a mobile.

In case this is too complicated, another option already mentioned in the analysis of the call tools is possible. If it is possible to share the screen and controls, the option of running the prototype on an own device and letting the user operate on it, sharing the screen with his mobile, is valuable.

Figure 14: Sketch logo

Due to the wide range of design applications, the study will focus on the most popular ones. According to the 2020 Design Tools Survey [51], the most popular applications for both design and prototyping are Sketch, Figma, Adobe XD, and InVision. It is also included in the Balsamiq study for its unique character and its high results in certain parts of the design process.

Based on the aforementioned study, the Maze study is also included as a tool to develop and automate usability tests.

Sketch: Vector graphics editor of Sketch B.V.

It is one of the first prototyping software and this has helped its popularization.

Sketch requires a license, and its dedicated app is limited to macOS. Additionally, its browser version allows accessing and viewing projects on other operating systems.

It integrates very elegantly simultaneous use for multiple users. In this way, it is possible for multiple designers to view the same project, work on it, and even collaborate on the design.

While it is not possible to run a prototype on mobile directly, there are certain additional applications that can allow this. This is the case with Sketch Mirror for iOS and Crystal for Android. Both tools allow the visualization of a prototype executed on a computer from the mobile device. Although the interaction is possible, it needs synchronization with the computer and, the interaction can be rough as the prototype is not running on the device itself.

Another existing option is Sketch Cloud. This tool allows you to view a prototype from a browser. However, the optimization of this is not ideal for mobile use.

Sketch's options to share and run the prototype on mobile exist but are more limiting. However, due to the popularity of Sketch, other tools are capable of importing projects designed in Sketch.

These other tools, more specialized in prototype sharing, can be a solution to not to discard the designing power of Sketch.

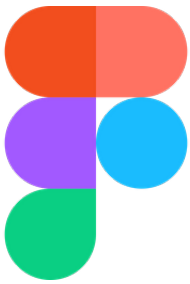

Figure 15: Figma logo

Figma: Vector graphics editor of Figma, Inc.

This fairly newer tool is very similar to Sketch. Figma is specially oriented to work in the browser. Additionally it also offers a desktop application. In both cases, Figma does not limit its use to one operating system.

Figma offers its service for free. It also offers different premium packages to facilitate its use in large projects with many collaborators.

Multiple collaboration by multiple people is allowed in Figma. It also pays special attention to being able to collaborate and comment live during the process.

Figma also needs a dedicated application to be able to run a prototype on mobile. This app is Figma Mirror. In the same way as the Sketch options, Figma Mirror needs to execute the project on a computer and to visualize that simulated execution.

Interaction is possible but it is also crude and limiting. There are also some third-party applications that can run Figma projects due to his popularity. Some of these applications are also specialized in prototype sharing and can help fill this gap.

Figure 16: Adobe XD logo

In general, for the study of this project, we can say that Figma and Sketch offer practically the same. However, it should be noted that Figma does not require a premium license and doesn't limit usage to macOS.

Adobe XD (Adobe Experience Design): Vector-based user experience design tool of Adobe Systems.

Tool similar to those already mentioned. In this case, with great integration with other Adobe services. This is a positive aspect since it allows you to work directly with .psd images imported from Photo Shop.

XD allows you to open a prototype from the Adobe XD mobile application. Unlike the previous options, this application allows to run the prototype. Thanks to this, the prototype can be used smoothly.

In any case, it must be remembered that this requires the download of the dedicated application. The teams should give the necessary permissions to the user in order to access the file.

Adobe XD requires a license to use its application, accessible from all operating systems. There is also the option to access XD from Adobe's Creative Cloud application suite. This package can be very useful by including other very popular design applications such as PhotoShop, After Effects or Illustrator.

Unlike the previous options, the use of XD would solve the option to share the prototype. Still, the service is not freely accessible as with Figma.

InVision Studio: Designing platform of InVisionApp Inc.

InVision was also one of the first tools of its kind. Until recently, it was the most used worldwide [51].

Similar to the other applications, InVision enables collaborative work. It is done from the browser, so there are no limitations based on the operating system.

InVision offers a free but limited service. In order to carry out more than one project, it is necessary to acquire a license. In addition, this license removes the limitation of 5 collaborative users

Figure 17: InVision logo

A direct system for testing with users its also offered. This is also accessible from mobile. Although the project can be viewed from the browser, it requires downloading the application to run the prototype.

InVision also offers the option to import prototypes exported from other design applications such as Figma or Sketch. With them, it is possible to access the Tester Mode and perform the usability test if necessary.

This is of interest as it can be a solution to run prototypes during the procedure without limiting the design application.

Figure 18: Balsamiq logo

Balsamiq: Mockup tool of Balsamiq Studios, LLC.

Balsamiq is one of the most unique prototyping tools. Unlike others, it focus on low and medium fidelity prototyping.

Instead of offering a complex framework that allows for very defined designs, Balsamiq offers a wide variety of assets. These helps the development of the prototype.

Balsamiq offers various tools to facilitate the execution of the prototype to test it. Still, there is the option to export the project as an interactive PDF to run as a prototype. This option is quite useful but it does not guarantee the optimization of the view to the proportions of the different possible mobiles.

Balsamiq works from a functional application on Windows, macOS and Linux. This allows collaborative work in a similar way to the other tools mentioned.

While its use and simplicity is a great advantage for early-stage prototype design, this can be limiting for working with more polished prototypes. Likewise, its sharing system is clumsy, although it can become functional.

Maze: User Testing system of Maze.

This tool has grown in popularity very quickly. It is included in this study due to its predominant use according to the 2020 Design Tools Survey [51].

This tool allows, from the browser, the creation of a flow to conduct a usability test in an automated way. Using blocks of different types, a process can be elaborated that can include questions of different types. There is also the option to create user tasks blocks. These tasks load a prototype and allow to define the expected flow.

During all interaction with users, their responses and interactions are collected.

Due to the collection of all this information, Maze opens the possibility to perform usability tests remotely asynchronously. Although this is not what we want to elaborate the procedure, it is good to point out this possibility. This option can be of special interest for carrying out a not too deep test with a large number of users.

The most interesting aspect is that the prototype that is loaded can be imported from Adobe XD, Figma, inVision, Marvel and Sketch. This eliminates the possible problems to run on mobile without limiting the design framework.

Figure 19: Maze logo

The possibility of also integrating question blocks can make it possible to dispense with other tools to launch surveys and forms. Unifying this possibility together with Maze would make it easier for the user to avoid having to understand different applications. In addition, this option allows you to quickly link the results of the forms with the interaction carried out on the prototype.

Maze can be run on mobile and computer through the browser. In this way, no download or registration is necessary for the users.

The Maze system is free but limited. In order to run several projects, it's needed to purchase a license. Likewise, this license also eliminates the restriction of 10 blocks for the usability test.

4.2.3 Forms and Questionnaires

During the development of the usability test it is possible to have support software to obtain data. This can be useful both for posing questions to the user and for completing one of the various standardized questionnaires as forms in usability studies.

The use of this software may be optional. In any case, it is always possible to record the results manually during or after the call. However, these tools can not only facilitate data collection but also order and visualize it.

There is a wide variety of applications dedicated to creating online forms. In fact, you can easily program your own system for this task. However, to streamline the procedure as much as possible, we will take the most popular ones for study. Based on this,the tools to be analyzed will be Google Forms, Microsoft Forms and TypeForm [54].

If any of these tools are to be used, their use may go beyond filling in forms. One option may be to integrate these forms with the completion of legal documents or agreements. By having shown the user one of these applications, reusing it for other reasons can avoid the user having to understand a different application. This will avoid confusion and facilitate the development of the procedure.

All of them allow easy sharing through a link to open from the browser. This means that no download is necessary.

Google Forms: Online form builder of Google.

This tool allows you to easily create forms from the browser. Forms offers a wide variety of different types of questions.

Forms also offers the option of reusing questions, even on different forms. This makes it the fastest option for some [54].

Thanks to its integration with the Google environment, it allows the exportation and visualization of data through Google Sheets. In addition, it offers very fast and simple visualizations of the users in the form of graphs and piecharts.

Figure 20: Google Forms logo

Another interesting aspect of Google Forms is its add-ons. The system allows the integration of functionalities developed by third parties. This opens the possibility of customizing the form. Thus, it is possible to solve more complex and specific needs.

Using Google Forms is completely free

Figure 21: Microsoft Forms logo

Microsoft Forms: Online form builder of Microsoft.

System very similar to Google Forms. It also allows you to easily create forms from a browser. Microsoft offers a smaller range of options for questions.

Being integrated with the Microsoft environment, this includes integration with Microsoft Excel. Similar to Google Forms with Google Sheets, data export and visualization is very easy. However, Microsoft Excel offers more alternatives for some of the question options.

Microsoft Forms is also free to use.

TypeForm: Online form builder of Typeform.

Although their functionality is similar, Typeform uses a different approach than Google Forms and Microsoft Forms. Instead of displaying a list of questions, Typeform only displays one question at a time. Typeform drives the entire form in a much more interactive and visual way. This can be positive and in it allows more information to be obtained.

However, this can also slow down the procedure and make it more tedious on long forms. This can also be a bit more complicated when reviewing the answers given.

A positive aspect of Typeform is that its more visual character makes it especially optimal for use on tablets and mobile phones.

Although it does not have direct integration with a data sorting application such as Microsoft Excel or Google Sheets, it allows you to easily redirect the results to other third-party applications (such as Slack) to work with the data from there.

4.3 Prototype

As explained in the introduction, the need for this study arises from the development of a payment application for shopping in stores. The idea for this application emerged as an innovation project. This project serves as a proof of concept for the formalization of a single payment application for the Spanish payment service provider Bizum. Currently Bizum does not have an offering in payment applications, with Apple Pay and Google Pay being the most used options in Spain [55].

Figure 22: Typeform logo

As has also been explained in the State of the Art, Bizum allows transfers between Spanish banks from the respective applications of those banks.

This means that the use of this technology depends entirely on the implementation and design of each bank. In addition, banking apps usually offer a wide variety of services today. This implies that the possible introduction of a new payment service in stores only makes the interaction within the banking application more obscure. Finally, this option would also provide a clearer and more defined corporate image.

4.3.1 System

To guarantee the payment service in stores it is necessary to define a complete system. The system is made up of three main elements as can be seen in Figure 23: The user's mobile device, the server system and the store payment terminal.

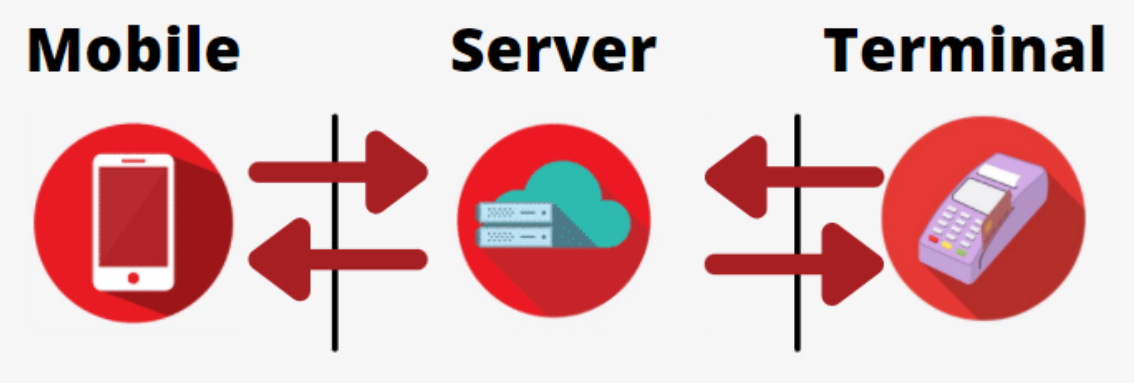

Figure 23: Components involved in the payment transaction flows.

Mobile Application: This is our payment application. Once the user is identified in the application, the user can manage payments from it.

Initially, the user must establish the payment method that it's wanted to use to communicate with the store's payment terminal.

Once communication has been established with the payment terminal and with the intermediary server, the payment information will be presented to the user. The user can thus accept or deny said payment. Additionally, the user is given the option to review all previous payments already made.

The development of this application has been carried out in parallel with the design of the interfaces, which are the ones that will be used in the so-called prototype.

The development of the mobile backend has been made in Java. The interface prototype has been made in Figma following the analysis made in the state of the art.

Payment terminal Software: This software is the one that contains the program to be executed by the payment terminals that want to allow payments through the proposed system.

Similar to the mobile application, it allows you to choose the type of payment through which to carry out the operation and allows communication with the central server.

Due to its smaller capacity and the objectives of the project, the interface of this software is outside the scope of this work.

The development of this software has been made in C language.

Central Intermediary Server: It is the server in charge of managing all payments and storing a record of them. Every time a transaction is initiated by a mobile phone or a payment terminal, the server is responsible for creating an entry in the database. As the transaction progresses, the server fills in that information with the data received from the mobile and the payment terminal. Finally, it offers all the information to the mobile so that the user is the one who decides whether to accept or reject the operation.

Once the operation is accepted, the server itself would be the one who would communicate with the Bizum payment system to complete the transaction, but this part is outside the scope of the proof of concept.

After simulating the payment, the server confirms to both devices the success or failure of the operation.

The flow followed varies slightly depending on the payment method selected. These flows are detailed in the next section.

The development of this back end has been made in Spring and using SQL for the database.

4.3.2 Flows

Different flows are established for the different payment methods.

The transaction is created in the server's database and, in order to be processed, it needs the information of the stores (such as address, payment concept or telephone number), the price of the payment and the user's information (user's telephone number).

When the transaction starts, an entry will be created in the database. In the case of QR generated by the payment terminal and SMS, it is the payment terminal that creates the transaction, attaching the information of the store and the price of the payment. In the cases of QR generated by the mobile and NFC, it is the mobile that creates the empty transaction and waits for a response for a while. User information is always added when the user accepts payment as a confidentiality measure.

Once the transaction is created, the payment method determines how to inform the other device of the ID of said transaction. When the other device receives this ID it completes the payment information and the user is offered the option to accept the payment. It is after this confirmation that the user information is shared.

Finally, the server is in charge of managing the payment. In this proof of concept the payment is simulated. Finally, both devices are informed of the result of the operation.

Each flow is individually detailed below and explained in a figure. Continuous lines indicate automatic actions. Dashed lines indicate manually performed actions.

QR code generated in the payment terminal: In this flow, the transaction is created by the payment terminal by choosing this option. Doing so also adds all the information of the trade and the value of the transaction. In response, the payment terminal receives the transaction ID in the server's database.

Once created, the payment terminal generates a QR code with the transaction ID.

Next, the user will read the QR code with the mobile, so that the mobile will be able to access to the transaction in the server and read all the information of the operation and display it. The user will be able to accept or reject the payment.

If the user accepts, the mobile sends the signal to the server to make the payment. This includes also the user information needed. Once the payment is completed, both the mobile and the payment terminal are informed of the success or failure of the operation.

This flow is illustrated in Figure 24.

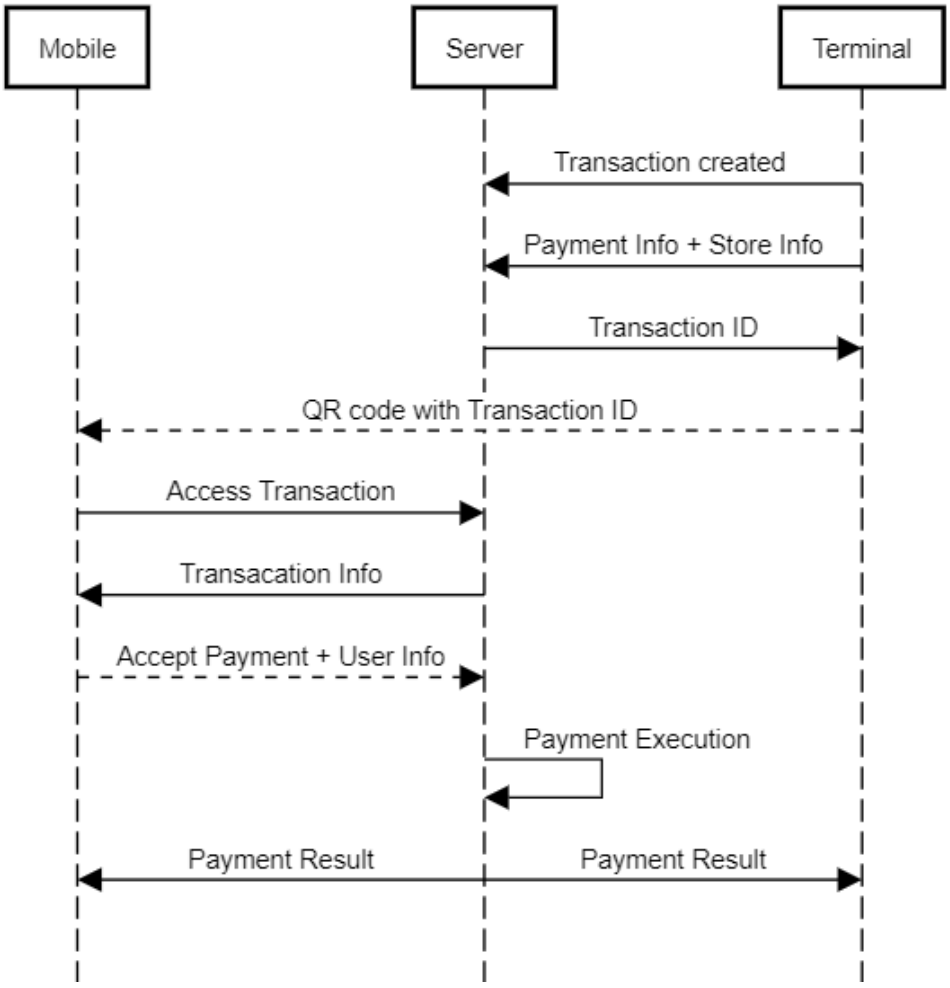

Figure 24: Payment flow for QR code generated by the terminal payments

QR code generated on the Mobile: In this flow, the transaction is created by the mobile by choosing this option. User information is not sended to the server yet. In response, the mobile receives the transaction ID in the server's database.

Once created, the mobile generates a QR code with the transaction ID and will wait for a response from the server for a set time.

Next, the clerk will read the QR code with the payment terminal. After this, the payment terminal will send and add the information of the trade and the value of the transaction. Then all the information of the operation will be sent to the mobile and display it. The user will be able to accept or reject the payment.

If the user accepts, the mobile sends the signal to the server to make the payment. This includes also the user information needed. Once the payment is completed, both the mobile and the payment terminal are informed of the success or failure of the operation.

This flow is illustrated in Figure 25.

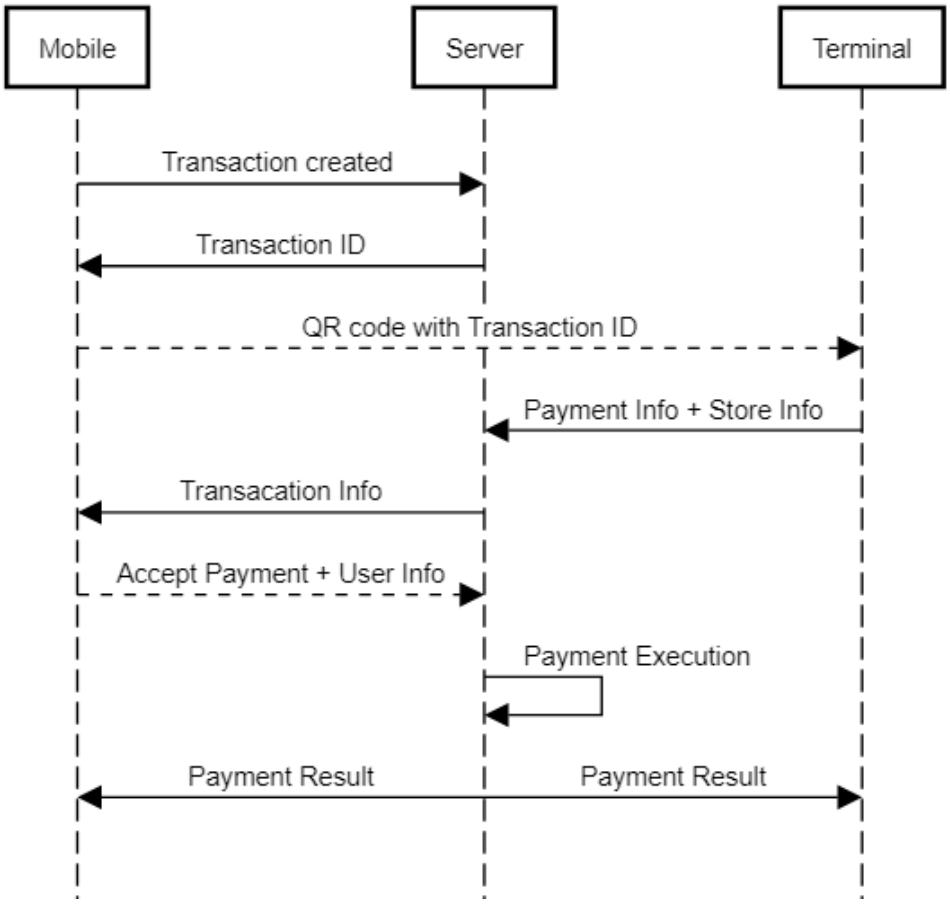

Figure 25: Payment flow for QR code generated by the mobile

NFC (Near-Field Communication): In this flow, the transaction is created by the mobile by choosing this option. User information is not sended to the server yet. In response, the mobile receives the transaction ID in the server's database.

Once created, the mobile sends an NFC message to the payment terminal with the transaction ID and will wait for a response from the server for a set time.

After this, the payment terminal will send and add the information of the trade and the value of the transaction. Then all the information of the operation will be sent to the mobile and display it. The user will be able to accept or reject the payment.

If the user accepts, the mobile sends the signal to the server to make the payment. This includes also the user information needed. Once the payment is completed, both the mobile and the payment terminal are informed of the success or failure of the operation.

This flow is illustrated in Figure 26.

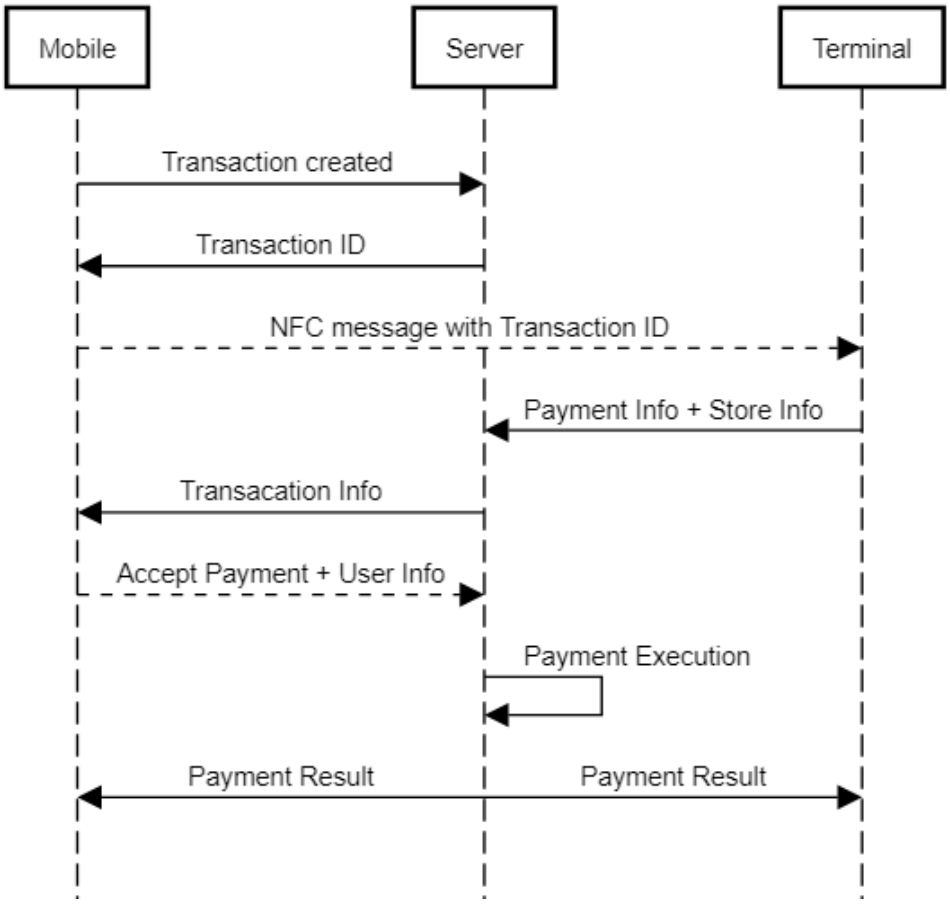

Figure 26: Payment flow for NFC payments

SMS: In this flow, the transaction is created by the payment terminal by choosing this option. Doing so also adds all the information of the trade and the value of the transaction. In response, the payment terminal receives the transaction ID in the server's database.

Once created, the payment terminal send a SMS to the mobile with a crypted code with the transaction ID.

Next, the user will introduce the code in the app, so that the mobile will be able to access to the transaction in the server and read all the information of the operation and display it. The user will be able to accept or reject the payment.

If the user accepts, the mobile sends the signal to the server to make the payment. This includes also the user information needed. Once the payment is completed, both the mobile and the payment terminal are informed of the success or failure of the operation.

This flow is illustrated in Figure 27.

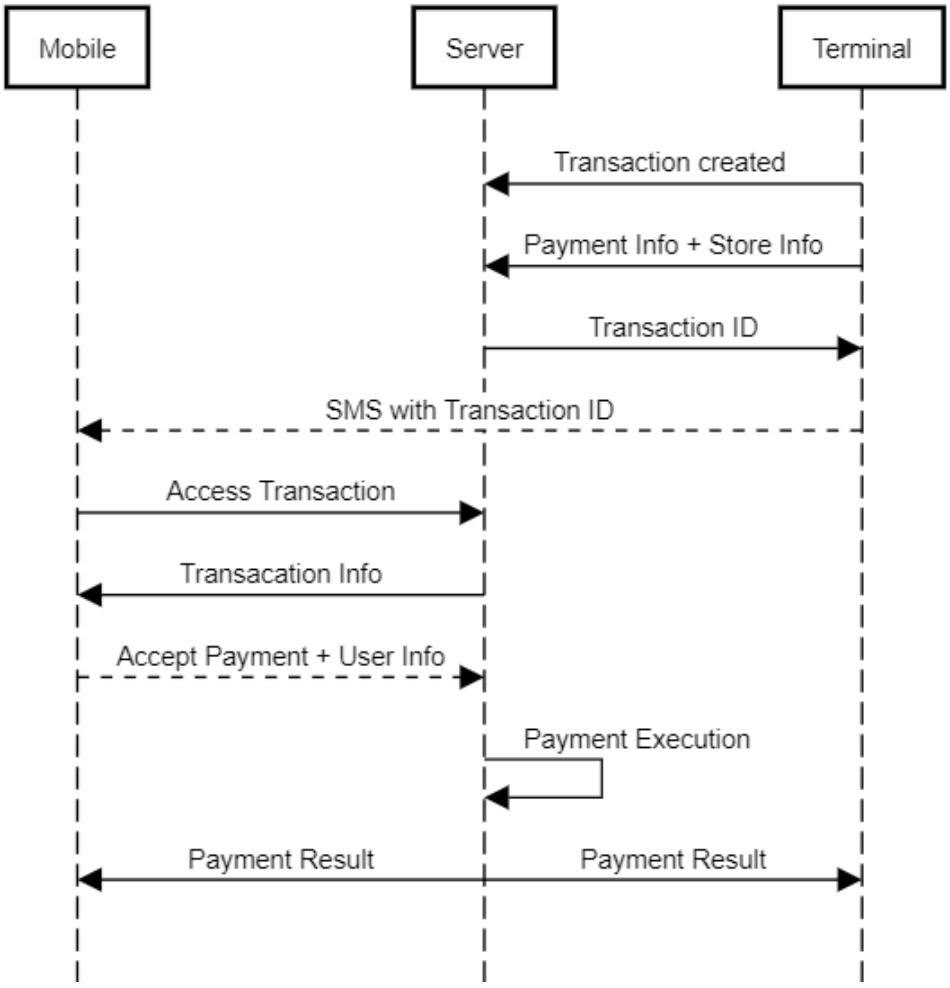

Figure 27: Payment flow for SMS payments

4.3.3 User Interface

For the necessary approach to decide the needed interface screens it is necessary to remember the objectives of the application. In this first proof of concept, the aim is for the user to be able to identify himself, choose the payment method among the aforementioned, make said payment and review the latest transactions.

Deliberately, for this part of the project, it has been left out of the scope to link the user's profile with their bank details, to register in the system and to manage the user's profile.

All screens have been designed and implemented in a prototype using Figma. On this prototype, the interaction to navigate from one screen to another has also been defined.

Launch: Screen displayed when the application is started. It is shown as long as the application finishes starting. Displayed in Figure 28.

Login: Screen that allows the user to choose the method of identifying himself in the application. In this phase, the options allowed for registering are fingerprint reading or entering the user's mobile number. Displayed in Figure 29.

Due to the early stage of the project, fingerprint access is not implemented. Access by entering the mobile number does not perform any actual verification and allow direct access to the main menu.

In later phases of the project it will be positive to devote special care to this part. An important aspect to decide is whether the account should be linked to a mobile number or to the mobile device. Multiple accounts from the same device can be allowed. It can also be possible to manage the same account from different devices.

Figure 28: Launch Screen

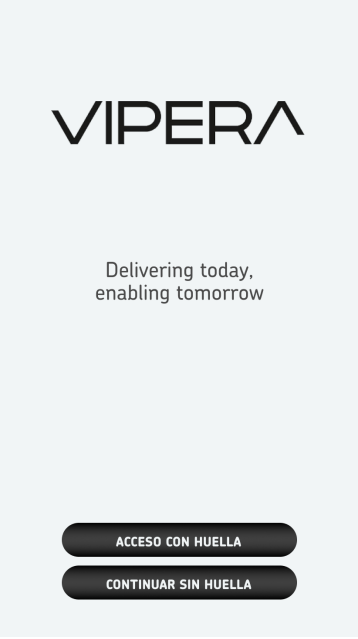

Figure 29: Login Screen

Main Menu: Screen that allows the users to choose the method of identifying themselves in the application. In this phase, the options allowed for registering are fingerprint reading or entering the user's mobile number. Displayed in Figure 30.

Main screen that shows the types of payment and allows to review the transactions made. This page is reached after the user has been identified. It is also returned to after successfully completing or declining a transaction.

In the future, should include access to the user's profile. Depending on the decision made regarding the use of multiple accounts, it may be appropriate to add the option to switch between accounts here.

Transactions: Also called "My Transactions". This screen shows previously performed transactions ordered in reverse order by age. In the list, separated by days, you can see the basic information of each payment. When clicking on any of the transactions, the detailed payment info screen is displayed. In Figure 31 can be seen this screen with a full list of payments done.

Currently the information stored and displayed is the most basic as a placeholder.

Detailed Payment: Screen that shows the complete information of a transaction carried out. Displayed in Figure 32 with full information of a specific payment.

Currently the information stored and displayed is the most basic as a placeholder.

Figure 30: Main Menu with payment methods

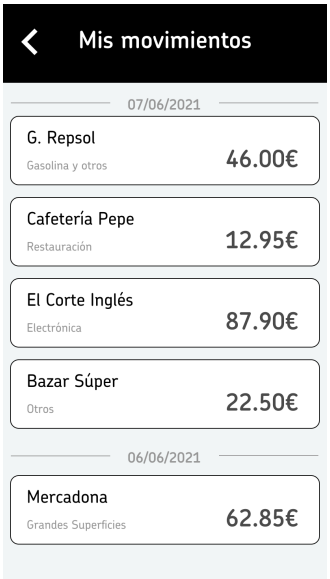

Figure 31: "My Transactions" list of payments

Figure 32: "Detailed Payment" of a payment

Generate QR: Screen that shows the QR code generated so that the payment terminal can access the transaction ID. Basic instructions are also provided to indicate that the code should be displayed to the payment terminal. In Figure 33 can be seen this screen with a QR containing the ID of a transaction.

Read QR: Overlay interface that is displayed on the rear camera of the mobile. It shows basic help to indicate that the QR code generated by the payment terminal should be read. To facilitate interaction, a guide area is also displayed to facilitate the code reading. This guide area is displayed as an overlay in Figures 34 and 35.

NFC: Screen with the instructions to make the payment with NFC. This screen will remain until the user brings the mobile closer to the payment terminal to allow the connection. Displayed in Figure 36.

Figure 33: Generate QR screen with a transaction ID coded

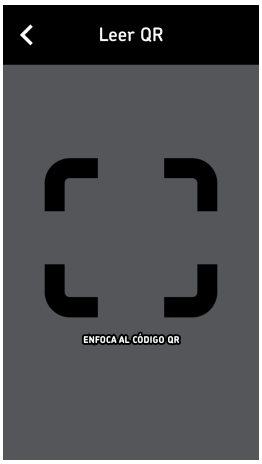

Figure 34: Overlay to be displayed when reading a QR core

Figure 35: Camera view with overlay interface for simulation

Figure 36: NFC Screen with minor payment instructions

SMS: The SMS payment flow begins on the Enter SMS Code screen. In this screen, the user must enter the SMS code received after the payment terminal entered the phone number. Displayed in Figure 37

In more polished versions, the option should be given to directly load the SMS code from that screen upon receipt.

Another option is to configure the SMS to directly open the Payment Confirmation screen. Since the SMS permits to obtain the transaction ID, that screen can be displayed with the complete payment information.

Payment Confirmation: Screen that is shown to the user with all the information related to the payment. The user is given the option to Accept and Cancel the transaction. Displayed in Figure 38.

In case of Acceptance, the user's information is shared and the payment transaction is carried out. If this operation is successful, the Accepted Payment screen is displayed. In the prototype the operation is simulated and, if the payment is accepted, it is taken directly to the Accepted Payment screen.

In case of rejecting the operation, the Canceled Payment screen is displayed. This screen would also be displayed in case the operation failed. In both cases, the information would explain the situation that occurred.

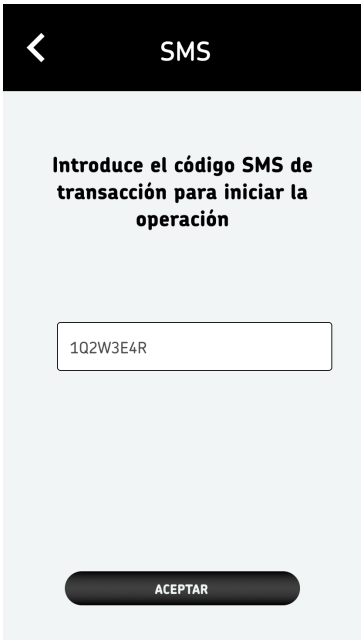

Figure 37: Screen for entering the transaction SMS code

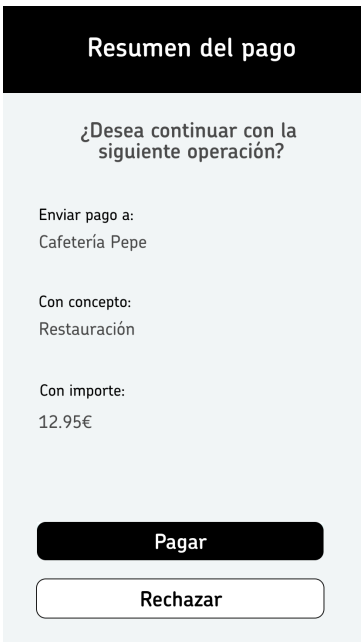

Figure 38: Payment Confirmation for a specific payment

Figure 39: Accepted Payment Screen

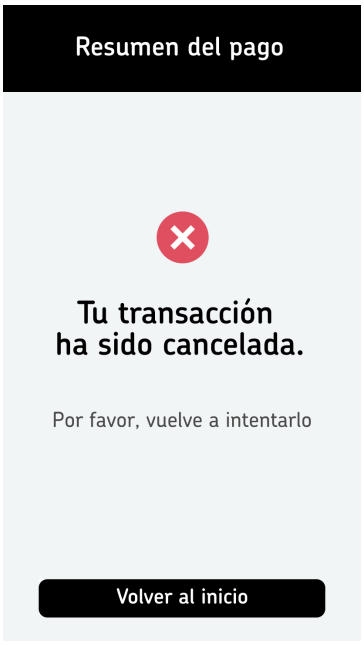

Figure 40: Canceled Payment Screen

Both the Accepted Payment and Canceled Payment screens allow the user to return to the main Menu. This screens are displayed in Figures 39 and 40 respectively.

4.3.4 User tasks

During the remote usability test, a series of scenarios are proposed to the user with a task to complete with the prototype. The objective of these tasks is to force the user to deal with the different functionalities of the application. This is to detect complications in the interaction. These tasks should simulate the actual use of the application.

For each task, a short introduction is provided explaining the context of the task. The goal that the user must reach is also indicated. These instructions are provided in writing, so that their wording is always the same in each test. A different wording or explanation may lead to a bias of the users. This, in turn, can affect negatively the data collection.

5 tasks are defined. In order to carry out a more exhaustive usability test, it might be positive to implement more functionalities to the prototype. This would allow the formulation of more complex and complete scenarios. In any case, since the final objective of the project is the evaluation of the procedure, the simplicity of the usability test will not have a great impact on the analysis of the project.

One task tests each different type of payment. A payment failure is included in one of the tasks. Additionally, the last task consists of the review of previous payments.

The usability test will change when the procedure is carried in different projects. Because of this, its definition is not part of the procedure design. This is why the

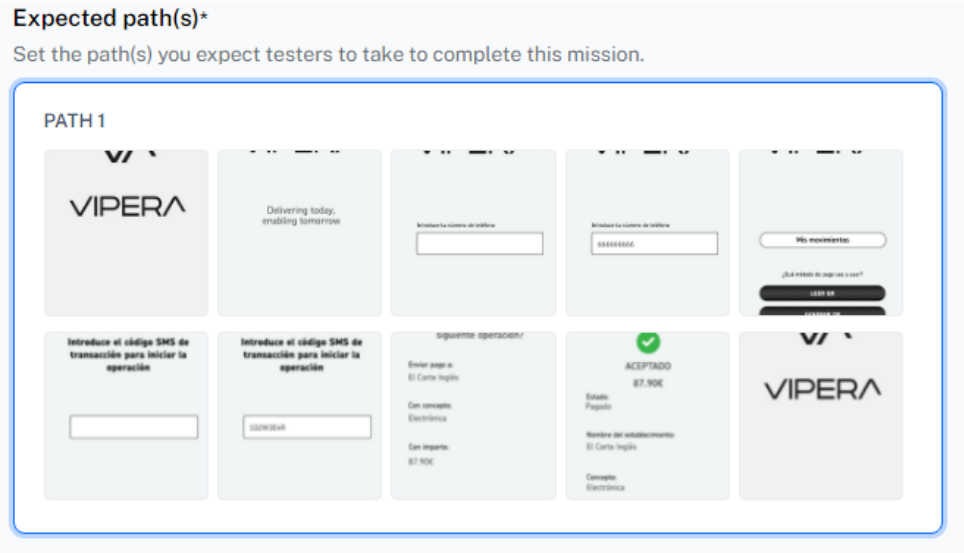

Figure 41: Expected flow for the first task set in the Maze project.

tasks remain the same in both iterations of the experiment. It is the difference in the results obtained that will be evaluated.

Tasks are builded as Maze blocks using the Figma prototype. A Maze project is created for the usability test. It contains a series of statistical questions together with the tasks explained below. Each of these tasks includes the scenario instructions and the flow that the user is expected to follow for its completion. One example of this tasks in Maze can be seen in Figure 41.

First Task: The first user task includes accessing the application from the Launch Screen. Fingerprint access is not enabled but this screen serves as a presentation of the access screen. It also makes it easy to ask about preferences in identification systems. The entry of the telephone number to access is simulated so as not to request it from the user.

The user is told to pay by SMS. You are informed that you have already given your phone number to the clerk and you already have the transaction code. From the main menu you must choose the payment by SMS and enter the code received. The entry of this code is also simulated.

After this, the user is shown the Payment Confirmation screen with the appropriate price and information as mentioned in the task statement. The user must accept the payment and the task is completed.

The wording of the task presented is as follows:

> You have gone to El Corte Inglés to buy a portable speaker and you are preparing to pay with the payment application.
> The clerk has asked you for your mobile number to send you the transaction code. You receive an SMS with the necessary code and you are ready to activate it.
> Complete the payment using the code received.

Second Task: This user task asks the user to generate a QR code so that it can be read by a waiter with the payment terminal. Due to the remote and simulated nature of the task, the reading by the payment terminal is simulated.

After this, the user is shown the Payment Confirmation screen with the appropriate price and information as mentioned in the task statement. The user must accept the payment and the task is completed.

The wording of the task presented is as follows:

> After completing your purchases at the Corte Inglés, you make a stop at Cafetería Pepe for a quick drink with a friend.
> When finished, you decide to invite him and pay through the payment application.
> The waiter brings you the payment terminal and offers to read a QR code that you generate.
> He completes the payment by showing a QR code.

Third Task: This user task asks the user to pay using NFC. This task includes an amount much higher than expected according to the task statement. Ideally the user will notice and cancel the payment. This task not only seeks to verify this type of payment but also to make the information given visible.

This task is conscientiously placed in third place. This is because, in order to seek the greatest possible realism in the interaction, the aim is that the user "gets comfortable" after carrying out the two previous tasks. With the more "relaxed" user, it is more possible that they do not pay an unrealistic major attention to the payment information due to the greater awareness of the user for being subject to study in comparison when making daily regular payments.

Due to the remote and simulated nature of the task, the NFC reading by the payment terminal is simulated. After this, the user is shown the Payment Confirmation screen with the inappropriate price. The user can accept or cancel the payment and in both cases the task will end.

The wording of the task presented is as follows:

> Returning home in your car, you stop at a gas station.
> You refill the entire deposit and you are ready to pay with the payment application through NFC (as if it were your own credit card), since you are in a bit of a hurry.

Fourth Task: This user task asks the user to make the previous payment by reading a QR code generated by the payment terminal. This time the payment amount will be as expected.

Once the option is selected, the QR read overlay interface will be displayed. This interface is displayed on a black background for a short time, simulating the loading time of the camera. Then, is a simulated image of the camera reading a QR code at the store is shown.

Due to the remote and simulated nature of the task, the reading of this QR code is simulated.

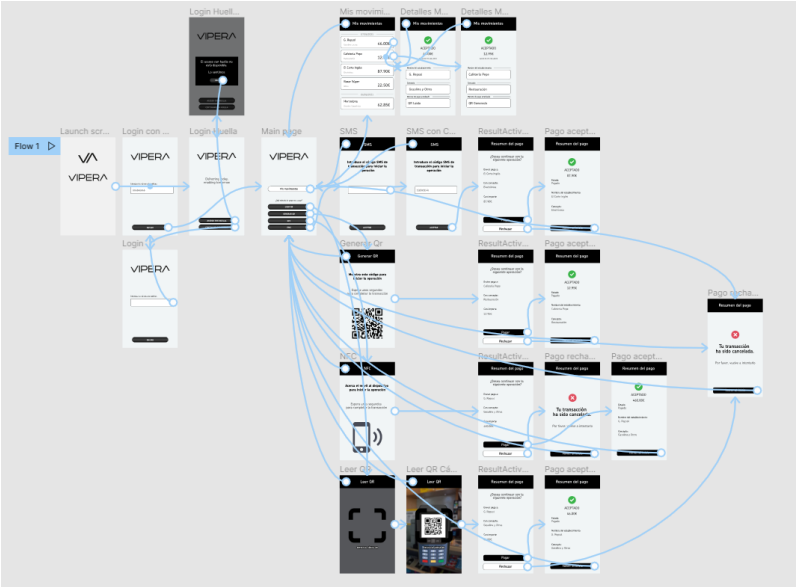

Figure 42: Figma prototype implementation with all screens and navegation used for the realization of the user tasks

After this, the user is shown the Payment Confirmation screen with the appropriate price and information as mentioned in the statement. The user must accept the payment and the task is completed.

The wording of the task presented is as follows:

> After verifying that the payment was incorrectly entered, the clerk restarts the transaction and shows you the payment terminal with a QR code.
> Complete the payment by reading this code.

Fifth Task: The user is asked to review the payment system used in the payment of the previous task. To do this, the user must access My Transactions where the payments made during the previous tasks will be shown along with some previous ones.

In this list, the user must click on the last payment, to access the Detailed Payment view, where the payment method is included. The task is completed after a short delay.

The wording of the task presented is as follows:

> To make sure that the payment at the gas station has been made correctly, check which payment method is the one that was finally used.

4.4 Users

In an effective usability test, it is possible to detect approximately 80% of usability problems with a sample of only 4 or 5 users [56].

However, it is necessary to remember the multi-layer character of this study.

Therefore, detect usability problems it's not sought but rather the efficiency of the procedure. This is why it is necessary to work with a larger sample, at least to evaluate the first version of the remote usability testing procedure.

This is why the sample of users should be wider than 5 users. It is necessary to evaluate the fluidity and results of the process in users with different profiles.

The use of the payment application prototype is a very useful tool for this. As it is a payment application, the user profile is also very broad. Its study also requires users with different profiles.

Therefore, the user sample must include users of different genders, ages, educational levels, economic fields and technological skills.

A priori, it is only expected that technological skills by itself is the one that can have a real impact on flow with the procedure. However, these skills often correlate with age.

The other criteria are not expected to have a large impact per se on the results. However, it is expected that they may have a lower impact to encourage or minimize the results obtained based on age and technological skills.

For the second iteration of the experiment, it is expected that a much smaller sample will be required. The size of the sample will depend on the results obtained in the first iteration.

In case the first experiment results in few failures being detected with the first proposed procedure, a smaller sample of users will be enough to validate the second procedure.

On the contrary, if the first procedure turns out to be very problematic, it will be necessary to make big changes. This will imply the need for a sample that can be as large as in the first experiment.

Additionally, conducting the first experiment will allow us to better adjust the sample size for the second iteration. Regardless of the good or results, it is possible that the original sample has not been large enough to obtain the expected results.

As long as the sample of users is diverse, the number of interviews can be adapted throughout both iterations according to the need of the experiment. It is always possible to increase the number of interviews if the results are poor. Similarly, if the are very similar and don't bring new information, it is possible to reduce the sample size.

With all this, we start from the idea of having a first iteration of between 8 and 12 users. The second iteration is expected to have between 3 and 6 users.

4.4.1 Possible bias:

It is worth highlighting a possible peculiarity that may bias the study. This is because the application initially seeks to work only in Spain (due to the scope of Bizum). Although the sample of users must include users with different profiles, they must live in Spain.

Presumably, nationality and culture do not have a great impact on the effectiveness of the procedure. However, it is not ruled out that cultural bias can have a small

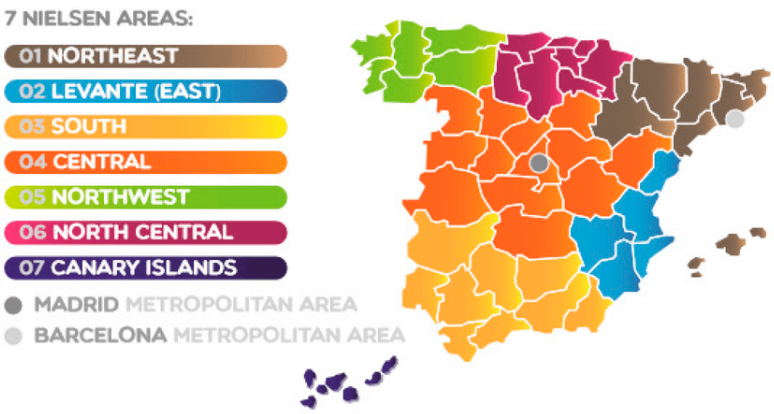

Figure 43: Division of Spain in the so-called Nielsen Areas [60]

impact. This is not only limited to the familiarity and preference of the applications used. It can also affect the predisposition and opinion of users to the procedure.

To guarantee the representative sample of the study and to reduce this bias as much as possible, location is added as an additional criterion for searching for users.

Therefore, users who live in Spain but spread throughout the country will be used. To facilitate this national representation, the Nielsen regions are taken as reference. This reference division of the Spanish territory, established by the commercial research company AC Nielsen, is frequently used to organize statistic, economic or market studies.

These regions are used in studies by universities [57], companies [58] and even in governmental studies [59].

Additionally, these regions are usually linked to a quota of representativeness. The objective is to ensure that the study is nationally representative, minimizing regional biases. This quota is adapted depending on the study's nature.

In any case, as these are small user samples, an attempt is made to respect these Nielsen areas but it is necessary to be flexible with the quota.

5 Solution and Results

This section shows the results of the experimentation. For this purpose, the two iterations performed are presented.

All pre-planning is explained. This includes the selection of tools and methods to be used. After this, the remote usability testing process to be evaluated is proposed.

For each iteration, the qualitative and quantitative results obtained from both the prototype and the procedure are presented. This allows to obtain a series of conclusions to be taken into account.

Finally, the final proposal of the validated remote usability testing procedure is presented.

5.1 First Iteration

The first iteration of the experimentation starts only from the knowledge acquired after the detailed study in previous sections.

In this first iteration all the tools to be used are selected. The basic structure for the procedure is also defined.

The results obtained allow to check the validity of the decisions made. They also allow to locate the pain points to be improved in the next iteration.

5.1.1 Planning

In order to propose the procedure, it is necessary to decide on the tools to be used and the steps to be followed.

5.1.1.1 Invitation and Call Preparation

The first task in the procedure is to be able to communicate with the user. This way, the usercan be summoned for the interview and given the information needed for the interview.

For this, an invitation message is defined that is common to all interviewees. This message contains the aforementioned information and links.

Seeking maximum flexibility, it was decided not to set a specific communication tool before the call. In this way, following the flow of conversations held with users to request their help, WhatsApp is mainly used to share this message. In one of the cases, email is also used.

For the establishment of the call it is decided to finally use Zoom.

If we take into account Zoom's premium service, it offers several advantages over Hangouts. On the other hand, Hangouts offers the option of calling without limitations for free.

Teams offers everything that Zoom can offer without the time limitation within its free package. However, creating a dedicated platform, while useful for the team, could overwhelm the user. Still, for large projects with recurring test teams, it is possible that it could prove to be of value.

Zoom is the only service that does not require registration by the user. Although, on mobile users must download the application (for which they are going to be guided), it does not require additional steps or complications.

For interviews of more than 40 minutes Zoom premium access is necessary. It is to be expected that interviews will require more than this amount of time on many projects.

In order to avoid this, it would be advisable to delay the mobile device to join the call. One way to do this is to ask the questions not related to the interaction before calling the user mobile. In this way, it is possible to extend the duration of the call, since the 40-minute countdown begins the moment a third member joins.

In any case, if the call is cut due to lack of time, it is perfectly possible to restart the call and continue. In fact, if it is no longer necessary to work with the mobile, this second call would not have a time limit.

It is decided to implement informed consent through Google Form. Using this tool is a way to unify form management. Later on, the SUS questionnaire will be carried out using Google Form and doing it on an application already used previously can make the experience easier for the user.

The use of Google Forms is not stipulated in the procedure as mandatory. If desired, it is possible to use another similar tool.

The invitation message is sent after confirming an appointment with the user to be interviewed. The standardized message for the invitation explains my personal situation with the project and includes the date of the interview, the link to the Zoom call and the link to the informed consent implemented in Google Form.

5.1.1.2 Call Setup

The moderator must initiate the call. Once the user joins, the moderator can help to activate the user's camera if it's needed.

After this, the user is asked to confirm if the informed agreement has been signed. If not, the moderator gives the option to help the user to do it. The user is also offered to solve any doubts that may have.

After confirming user's agreement, recording of the session can be started using the built-in Zoom option.

The user is then introduced to the project. It is necessary to explain him the process to be followed and its objectives. It is also important to help relax the user by reminding him that there is not need to worry about the answers given.

5.1.1.3 Test Preparation

The user can be asked pertinent questions prior to the interaction. These can be demographic, about their experiences, tastes, etc. In the case of this study, these are limited to demographic data.

It is decided to use Maze for the sharing of the usability test. The great advantage of Maze is its ability to create a flow that includes the defined user tasks themselves. This is not only very useful to be able to proceed from one task to another auto-

matically, but it also avoids the need to download or configure more applications or projects by the user.

Through a single link, the user can easily perform the task from their mobile. This link is deliberately not included in the invitation email to avoid the possibility that users will execute it when they should not.

Since Maze therefore runs on the user's mobile, it is necessary to integrate this into the call. Maze records user inputs, so a post-interview study of user interaction is possible. Even so, it's useful to be able to follow the user's actions in real time. Therefore, it is time to invite the mobile to the call to be able to share the screen.

To do this, another standardized message is sent to the user with the access link to the Zoom call. To do this, the previously used communication channel can be reused. It is also possible for the user to access the call from the link sent in the invitation message already sent before the meeting.

At this point, if the user does not have Zoom, which is to be expected, they must download the application. For this it can be do from the store corresponding to the mobile's operating system. However, it should also be given the option to access the store's download screen from the call access link itself.

Although the user must carry out this task, it is possible to give support when the user is already on call. The user can be advised and may even be asked to show your phone through the camera if complications arise. This last option, although crude, can be useful for the most unforeseen cases.

Once the user has Zoom on their mobile and joins the call, the useris asked to share the screen in Zoom.

5.1.1.4 Usability Test

Once the screen is being shared from the mobile, a new standardized message is sent with the link to the Maze test. Once again, the communication channels that are already in use are used. It is also possible to share it from the Zoom chat, but this is valued as something more confusing than email or WhatsApp.

Once the user opens the Maze link, the usercan start their own usability test. Although the user can do it alone, it is advisable to guide him through this process.

In this case, and as a test, several demographic question blocks have been included. An information block has also been included to clarify how to simulate the interaction with the payment terminal in the user tasks due to being operating remotely. This block is also used to explain the importance of the think aloud protocol.

Once the maze test and all the user tasks have been completed, a discussion is held with the user about the experience during the interaction.

It is advisable to remove the user's mobile phone from the call at this time. This prevents him from being distracted. In addition, this way the mobile is available as a support tool in case of wanting to share additional documents for the conversation.

For the evaluation of the experience in the process, the SUS questionnaire was used. This not only allows evaluating the user's opinion but also serves as a comparative tool in the two iterations of the experiment.

To fill out this form, it is decided to implement it through Google Forms. However,

instead of sharing it with the user, it is the moderator who shares their screen and fills it in by the user. Once the user is viewing the form, it is explained what it consists of and is asked to answer. It is the moderator who enters the input, but the user sees his answers and all the questions.

This seeks to make this part as easy as possible for the user. With this solution, the user does not have to execute anything additional. In addition, by being on the call and the moderator being the guide through the questions, the user has the possibility of asking easily if they do not understand something.

The tool used for the form is not strictly stipulated in the procedure. Since it is the moderator who runs the form, any option is valid since it will not pose an added difficulty for the user.

In any case, it is recommended that this tool be unique in case of using several forms.

Once the prepared questionnaires are completed, the discussion can be continued with a fluent conversation. When all the questions have been asked and the moderator considers that all the necessary information has been obtained, the procedure can be closed.

5.1.1.5 Procedure Evaluation

In the case of the experiment, the process continues to evaluate the procedure. This includes the completion of the NASA-TLX questionnaire following the same dynamics as with the SUS questionnaire. There is also a discussion about the procedure itself. The moderator say goodbay to the user and the meeting ends.

5.1.2 Procedure Proposal

In order to carry out the procedure properly, the moderator must have the following:

- Informed Consent and Legal Documents. Coded in the **application the research team prefers to use**.
- **Zoom** room
- **Maze** test with the integrated Prototype. This prototype should be imported from **Adobe XD**, **Figma**, **inVision**, **Marvel** or **Sketch**.
- Forms (if necessary). Coded in the **application the team wants**.
- Other possible elements that the team considers necessary for its usability test.

The procedure is defined in the following steps.

1. **Invitation:** The user is summoned to a meeting on the agreed date. For this, a standardized message is sent to the user by **the available communication method**. This message should include:
    - **Zoom** call link and access instructions.

- Link to informed consent or other legal documents. The tool to use is at the discretion of the team.
- The devices the user will need. Presumably a computer or tablet and a mobile.

2. **Introduction:** Call starts in **Zoom**.

   The user is asked to confirm that they have signed the informed consent and the pertinent legal documents. If the user has not done it, the moderator will give instructions and support to do so.

   Once the documents have been signed and confirmed, the session can begin to be recorded.

   Next, the project that is being carried out is explained in detail. The process to be followed and its objectives are also explained.

   If there are questions or forms to fill out prior to prototype run, they are done at this time.

3. **Setting up the mobile:** A new standardized message is sent through the communication channel established with the **Zoom** call link and access instructions. This link can also be open by the user from the meeting invitation message.

   If the user does not have Zoom on his mobile, the interview give him instructions to download it. Opening the access link, it will take the user to the corresponding store page to start the download. The download can be done also by searching for Zoom from the store.

   Once Zoom is available, the user can join the call from their mobile using the access link or the credentials sent.

   Once the mobile phone joins the call, the user is asked to share their mobile screen.

4. **Remote Usability Test:** A new standardized message is sent through the communication channel established with the **Maze** link.

   When the user open this link, the browser will open showing the test created.

   The moderator will guide the user through the Maze test. This process will change completely based on the needs and reality of the prototype to be evaluated.

   It is important to indicate to the user the importance and necessity of executing the think aloud protocol.

   Once the test in Maze is finished, the moderator can remove the user's mobile phone from the call.

   Next, if a questionnaire is to be carried out, the moderator will share their computer screen, showing the form to be completed. In case the user needs

support for an additional file, a new standardized message is sent through the communication channel established with the needed files.

The moderator will guide the user through the forms, being the one who completes the questions following the user's instructions.

The discussion and the questions that the moderator want to ask can be done while the user continue on the call.

5. **End and farewell:** Once the usability test is finished and it is considered that all the information has been obtained, the moderator can say goodbye to the user.

   If the moderator wishes, the recording can be stopped. In any case, the recording will stop automatically when the moderator close the call.

   The procedure ends here and the moderator can close the call. Next, the call recording will be processed to generate the files corresponding to the meeting.

5.1.3 Experimentation

Initially, two pilot tests were carried out. These allowed us to detect a couple of bugs related to the Maze configuration.

After this, this first iteration finally consisted of 10 experiments. These were held between June 14 and 21, 2021.

The ages of the users interviewed are between 25 and 65 years old. 6 of them are women and 4 men. At least one user is from each of the 7 Nielsen regions mentioned in the previous section. Also, 3 users belong to the metropolitan area of Madrid.

5.1.3.1 Usability Tests Results

To check the validity of the procedure, it is necessary to evaluate the results of the Usability Test. This data will also be useful for comparison with other iterations of the procedure.

The general opinion of the prototype was to highlight its simplicity. All users also said they were happy with the functionalities and their interaction.

Still, certain problems did appear. Some of them did it repeatedly. Severe failures can be considered to be those that occur recurrently.

Severe failures could also be those that have a major negative impact on task completion, but the problems reported were not from this category.

The most prominent failures so are the following:

"My transactions" button: A couple of users had trouble locating the "My Transactions" button. The button is displayed in white on a white background and this can make it difficult to locate. It would help to use a second color that stands out to distinguish it from the payment options without losing visibility.

Realization of the wrong payment: The intentionally erroneous payment in task 3, although it was detected by all users, 3 of them did it just after

accepting said payment. Users considered that the information displayed was sufficient, but due to the rush it is easy to confirm by mistake.

Half of the users (including the 3 users who failed) said they would value a second payment confirmation positively. Since the other half disagreed, perhaps the best option would be to offer a customizable setting to show double confirmation or not.

QR payments naming: The "Read QR" and "Generate QR" options together were a bit confusing for 3 users. They explained that they were used to reading QR codes but not generating them. One of the users came to think that generating the QR could be a form of reimbursement. Another user thought that he should read first in order to generate later.

Perhaps it would be positive to look for more explanatory terms or to accompany them with explanatory icons.

Favourite payment method: All users positively value having different payment methods. However, 4 of them rated as positive being able to establish a favorite payment method. In this way, it is not necessary to choose the type of payment each time.

Since not all users want the option to set the favorite payment, the best solution would be to include this feature in an optional and configurable way.

Regarding the SUS questionnaire, the result obtained was 87.25. This means according to the standards that can be seen in Figure 44 that the result obtained by the prototype is very good. Still, based on the comments received, there is room for improvement.

5.1.3.2 Procedure Results

The experiment was completed with satisfactory results. The process was unanimously evaluated positively. but some pain points were detected.

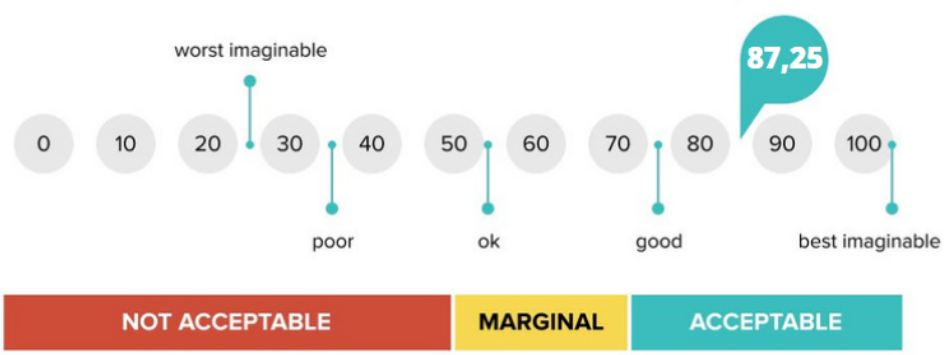

Figure 44: SUS score Acceptability Scale for first iteration result [44]

The most useful interviews were those conducted with older users. These users were those with the least technological skills and confidence, which helped to detect more possible problematic aspects.

On the contrary, interviews with more skilled users, especially younger ones, could be carried out without any complications.

The main pain points detected and non-conformities were the following:

Intrusiveness: 5 of the users expressed that sharing the mobile screen could be intrusive. Especially if they had to access a personal communication tool to open the Maze link. This opinion becomes more critical if the means of communication is WhatsApp instead of Email.

These users also commented that knowing this need in advance would have allowed them to prepare and mentalize. And in this way they believe that this problem would be avoided.

When to fill forms: The majority of users valued the way of filling in the forms very positively. However, two of them assured that they would have preferred to complete the form on their own.

These users stated that they could feel conditioned to answer while on call with the moderator.

One of them suggested pausing during the call to give him time to answer him. The other said that he would have preferred to do it at the end of the experiment.

Communication tool: 4 of the users expressed preferring WhatsApp as a form of communication due to its lightness. They said that even so, using the mail would not be a problem for them.

3 of the users said they prefer email because of its professionalism. They did express their misgivings about using WhatsApp. They would rather share their mail than their phone number. In addition, the need to share screen was less intrusive in email than WhatsApp.

When to download Zoom: Most users positively valued downloading Zoom on their mobile during the call. However, 2 of them said they preferred to know in advance that it was going to be necessary to have Zoom. Had they known, they would have downloaded it beforehand.

One of the users commented that it could be the case of not having enough memory space for the download, which would complicate the procedure.

Permissions for sharing screen on Mobile: When starting to share screen from Zoom, users' mobiles request permissions to be able to share screen and be able to do it while other applications are running.

How to configure these permissions depends on the mobile and operating system of each one. This problem was not expected to be detected, but as the

moderator was already on call, it was possible to give assistance to users who felt more confused. In this way, all users finally managed to share a screen.

5 of the users stated that this part was the most confusing of the procedure.

Additionally, another series of aspects to take into account were detected. Although they are not especially critical aspects, they are aspects to take into account.

Auto audio recording in Zoom: If the user has previously used Zoom on their mobile, it is likely that, when joining the call, the audio will be activated by default. Since the user is also on call from another device, this can create a bit of fuss and confusion when audio starts to play on both devices simultaneously.

Download Zoom link: When opening the call link from the mobile without having Zoom downloaded, the option is given to open from Zoom or from the store.

In some cases this selection has been confusing or has not worked. In such cases, it was recommended to directly open the store and search for Zoom from there.

User trust: One of the older users pointed out the importance of building trust in the invitation message and during the introduction.

Currently, there is a large number of scams aimed at taking advantage of the lack of technical knowledge of older people.

This user believed that it was important to avoid mistrust because of this.

Connection quality: Although the consumption of data necessary to develop the procedure is very low, this can be a limiting element. In the case of very poor coverage, the call may not be maintainable.

Also, the process of executing the Maze link requires a small load to work. If the coverage is very poor, it is possible that this load failure.

It is necessary to point out an aspect that generated some confusion in the experiment. The Nasa-TLX questionnaire was carried out in a similar way to the SUS questionnaire. The questionnaire was implemented in Google Forms. The user was shown sharing the moderator's screen and the moderator was answering the questions following the user's guidelines.

Although this method of filling forms was positively valued as part of the procedure, most users came to consider the Nasa-TLX questionnaire as part of the procedure itself. Due to being a complex form, many users answered that the Nasa-TLX questionnaire had been the most complex part of the procedure. Despite clarifying the scope of the procedure prior to that question, it was needed to be explained again.

The objective of trying to standardize the tools during the experiment itself, came to influence so it was not entirely clear when the procedure ended and the

Weighted		**Raw/Unweighted**	
Overall	32,07	**Overall**	32,50
Diagnostic Subscores		Diagnostic Subscores	
Mental	112,00	Mental	33,00
Physical	20,00	Physical	12,00
Temporal	165,00	Temporal	31,00
Performance	70,00	Performance	20,00
Effort	43,00	Effort	29,00
Frustration	135,00	Frustration	25,00

Table 5: NASA-TLX results for the first Experiment

experiment continued. Although it may seem counterintuitive, using a different tool would help clarify the paradigm shift.

The results of the Nasa-TLX questionnaire are shown in Table 5. Although these data do not have a direct reading, they allow us to detect the aspects that most affect workload.

These results can also be used in a comparative way. Apart from comparing them with those obtained in the second experiment, it is possible to compare them with other similar experiments.

Considering the procedure a "Cognitive Activity" as defined by Rebecca A. Grier [47]. She defined congnitive activites as *"Tasks requiring mental action such as computer programming, flight, planning, proof-reading, speech shadowing, etc... "*.

Comparing with her study, it can be said that the results obtained are less than the 25% percentile for this type of task (38.00). The results are also less than the 20% percentile of the global result.

It can also be seen that the values that contribute the most to the general load are the temporary load and the mental load. Additionally, the character with the least impact is the physical load. These results were to be expected given the nature of the procedure.

It is necessary to emphasize again that the interpretation of these data does not give direct information. However, with the comparative study we can induce that it is a mostly simple process.

In any case, despite this small complication, once it was re-explained to the users, the evaluation could continue successfully.

5.1.4 Conclussions

The results of the procedure were promising and exceed initial expectations. However, there are aspects to improve. This is why a new iteration of the experiment is required. This can be solved with a small sample of users given the good results.

It is especially necessary to study how to solve the possible suspicion of sharing the mobile screen. Form management, when to download Zoom, and the communication method to use should also be reviewed.

As positive aspects, the validity of Maze and Zoom as fixed tools for carrying out the procedure is proven.

5.2 Second Iteration

After obtaining the results during the first iteration it is possible to draw certain conclusions. The results show the validity of certain decisions taken, but also show existing and improvable problems in the procedure.

This second iteration seeks to improve the overall experience of the procedure while maintaining the quality of the information obtained.

The proposed changes in the process are validated with a new iteration of experimentation to finally validate the final remote usability testing procedure to be proposed.

5.2.1 Pain points and changes

Based on the results obtained in the first experiment, it was decided to maintain the use of Zoom and Maze for the procedure. In addition, 5 key pain points are detected in which to make improvements:

1. Intrusiveness of mobile screen sharing: The easiest way to fix this problem would be to skip the screen sharing step in Zoom. However, after the analysis carried out prior to the experiment, no such comfortable or simple alternatives were discovered. Therefore, this problem must be assumed for now.

Most of the users who considered that screen sharing could be a problem agreed that this would be mitigated by avoiding the use of WhatsApp and giving advance warning.

Informing in the invitation that the user will have to share their mobile screen is a simple solution and allows the user to be forewarned. In the case of not agreeing with this, it is more time-saving that the user refuses to start the interview than dropping once it has started.

Regarding the use of WhatsApp, this is also avoidable by using email as a communication method. This also affects the third pain point.

Although a majority of users said they prefer to use WhatsApp, none were reluctant to use the mail. Using email as a communication method also avoids having to request a phone number, which was also considered more private than email.

There is also the option of using the internal zoom chat to share the maze link once the call has been initiated and the call has been shared. Although this would prevent images of the email from being shared, it is considered that it will be easier for the user to open their email than having to learn to use the zoom chat when screen sharing is already taking place.

In addition, it is necessary to have sent a previous invitation that cannot be sent by Zoom. Having had to access a link from the email, how to do it will already be known to users.

Despite losing flexibility with respect to the first solution, it is considered that stipulating the mail as a communication tool can allow solving pain points 1 and 3.

Even so, in future experiments it could be evaluated to use the Zoom chat to share the Maze link when the screen is already being shared.

2. Method and timing of filling in forms: It is difficult to find a universal solution for this pain point.

It was considered offering whether to fill out the form during the call using screen sharing or to fill it out later. However, it is possible that doing this differently may have a higher bias in the result than it is intended to avoid.

Therefore, for this experiment it is decided to continue with the defined method. In any case, users will be asked if they would have preferred to be able to complete the form after the experiment.

3. Communication method before and during the call: This point is resolved after the resolution taken in the first pain point.

The first iteration concluded with a positive reception to using Whatsapp. In addition, the possibility of making the procedure more flexible in order to use the means of communication preferred by the user was positively valued.

However, standardizing email as a means of communication will allow, although losing flexibility, to avoid privacy issues without negatively affecting users who do not have a problem using other applications.

4. When to download Zoom: Regarding this pain point, it should be noted that the majority of users positively valued having had support for downloading Zoom. Even so, to avoid this step for more skilled users, it is decided to inform in the invitation that Zoom will be needed on the mobile to make the call. Additionally, the option of being able to perform this download during the call will be offered to have support.

This option should provide a solution to those who preferred to avoid downloading during the procedure without leaving those who preferred support unattended.

5. Give permission to share screen in Zoom: This problem is an especially tricky one to solve.

The settings to give permissions to the applications depend on each mobile and operating system. While certain similar cases were found, too many different shapes were discovered to provide universal guidance.

As an additional complication, screen sharing is not yet so to provide support only the computer or tablet camera and user explanations are available.

Perhaps it would be possible to carry out a comprehensive study of the most common systems and offer instructions with pictures so that the user can solve the problem. However, given the complexity of this, it is decided to discard this idea for this experiment.

Thus, this problem continues to be indicated as a pain point that currently has no choice but to assume.

After this analysis, the second iteration of the procedure is formulated. The changes made therefore are the following:

- Set email as a communication method before and during the call.
- Inform of the need to share screen in the invitation message.

- Inform about the need to have the Zoom application on the mobile. Still, provide the option to download during the interview.

To make it easier to read the new procedure, changes to the original formulation have been highlighted in blue.

5.2.2 Procedure Proposal

In order to carry out the procedure properly, the moderator must have the following:

- Informed Consent and Legal Documents. Coded in the **application the team wants**.
- **Zoom** room
- **Maze** test with the integrated Prototype. This prototype should be imported from **Adobe XD**, **Figma**, **inVision**, **Marvel** or **Sketch**.
- Forms (if necessary). Coded in the **application the team wants**.
- Other possible elements that the team considers necessary for its usability test.

The procedure is defined in the following steps.

1. **Invitation:** The user is summoned to a meeting on the agreed date. For this, a standardized message is sent to the user by **email**. This message should include:

   - **Zoom** call link and access instructions.
   - Link to informed consent or other legal documents. The tool to use is at the discretion of the team.
   - The devices the user will need. Presumably a computer or tablet and a mobile.
   - Soft request to download the **Zoom** app. The user will also be given the opportunity to download it during the call if preferred.
   - It must be announced that during the procedure the user is going to share their mobile screen.

2. **Introduction:** Call starts in **Zoom**.

   The user is asked to confirm that they have signed the informed consent and the pertinent legal documents. If the user has not done it, the moderator will give instructions and support to do so.

   Once the documents have been signed and confirmed, the session can begin to be recorded.

Next, the project that is being carried out is explained in detail. The process to be followed and its objectives are also explained.

If there are questions or forms to fill out prior to prototype run, they are done at this time.

3. **Setting up the mobile:** A new standardized message is sent through **email** with the **Zoom** call link and access instructions. This link can also be open by the user from the meeting invitation message.

   If the user does not have Zoom on his mobile or decided to not download it, the interview give him instructions to download it. Opening the access link, it will take the user to the corresponding store page to start the download. The download can be done also by searching for Zoom from the store.

   Once Zoom is available, the user can join the call from their mobile using the access link or the credentials sent.

   Once the mobile phone joins the call, the user is guided to share their mobile screen by clicking on the big green Share button and then selecting Screen.

4. **Remote Usability Test:** A new standardized message is sent through **email** with the **Maze** link.

   When the user open this link, the browser will open showing the test created.

   The moderator will guide the user through the Maze test. This process will change completely based on the needs and reality of the prototype to be evaluated.

   It is important to indicate to the user the importance and necessity of executing the think aloud protocol.

   Once the test in Maze is finished, the moderator can remove the user's mobile phone from the call.

   Next, if a questionnaire is to be carried out, the moderator will share their computer screen, showing the form to be completed. In case the user needs support for an additional file, a new standardized message is sent through the communication channel established with the needed files.

   The moderator will guide the user through the forms, being the one who completes the questions following the user's instructions.

   The discussion and the questions that the moderator want to ask can be done while the user continue on the call.

5. **End and farewell:** Once the usability test is finished and it is considered that all the information has been obtained, the moderator can say goodbye to the user.

   If the moderator wishes, the recording can be stopped. In any case, the recording will stop automatically when the moderator close the call.

   The procedure ends here and the moderator can close the call. Next, the call recording will be processed to generate the files corresponding to the meeting.

5.2.3 Experimentation

Initially, one pilot interview were carried out. This confirmed that the experiment could be carried out correctly.

After this, this second iteration finally consisted of 3 experiments. These were held between August 5 and 7, 2021.

The ages of the users interviewed are between 27 and 72 years old. 1 of them are women and 2 men. All of them were from different Nielsen regions.

5.2.3.1 Usability Tests Results

In this case, to evaluate the effectiveness of the new procedure, two things need to be evaluated.

On the one hand, the information obtained is useful. As already mentioned, if the usability test does not report useful information, the whole procedure is meaningless.

On the other hand, that the results obtained are similar to those obtained in the previous iteration. Very disparate results in both the opinions and the SUS score may imply a possible bias in the procedure.

Therefore, it is necessary to analyze the responses and comments received by users.

> **"My transactions" button** One of the users needed time to locate the "My Transactions" button. She claimed that the white button on a white background did not attract his attention enough.
>
> **Realization of the wrong payment:** One of the users accepted the wrong payment from task 3. Even so, he claimed to have realized it right after accepting.
>
> Both this user and another agreed that a double commit could avoid mistakes without becoming a tiring task.
>
> **QR payments naming:** One of the users commented that the "Read QR" and "Generate QR" buttons were confusing. Working in a hurry, she did not distinguish them properly.
>
> Once the confusion was cleared up she said she would be able to distinguish both options. In any case, using another phrasing might be more explanatory.
>
> **Favourite payment method:** Two of the users said they would prefer to have the option to set a favorite payment. In this way, they would not have to choose the type of payment for each occasion.
>
> In any case, all users positively valued having the various types of payment.

These comments are very similar to those obtained in the experiment. In addition, they are perfectly valid as a result of the usability test. In this way, the effectiveness of the procedure can be confirmed.

The SUS form results were a score of 86.66. This is slightly lower than the score of 87.25 from the first experiment. In any case, a variation of just 0.59 can be considered marginal.

These results show not only the good valoration of the prototype but also prove that the changes in the procedure have not produced any bias in the results.

Charts.

5.2.3.2 Procedure Results

Again, the evaluation of the experiment was very positive. All users agreed to mention the simplicity and ease of this.

Users were asked about the proposed changes to assess their opinion regarding them:

> **Communication method:** All users confirm that email seemed like a good communication system. As they were not given other prior options, they all believed that it is the best communication system.
>
> **Intrusiveness of mobile screen sharing:** All users said they were happy with screen sharing because they were already forewarned. One of them assured that, had he not known beforehand, he would surely have been a lot more uncomfortable.
>
> **When to download Zoom:** All users positively valued having been informed of the need to have the Zoom application on their mobile. It was not necessary to support any of them to do so.
>
> Two of them doubted they would have had a problem loading anything from the store. The third and oldest of them stated that, had they had problems, they would have waited for the meeting to get support.

Additionally, the pain point related to giving screen sharing permissions remained unresolved.

In this iteration, only one of the users had some complication with this aspect but again it was solvable thanks to being able to give support from the call.

Regarding the other minor comments noted during the first iteration, there were no new discoveries. It is necessary to know their existence, but knowing about them they are easily manageable before or during the call. It is easy for the moderator to gain more fluency and experience during the experiments and will be able to take these small details into account.

The results of the Nasa-TLX questionnaire are shown in Table 6. Although these data do not have a direct reading, they allow us to detect the aspects that most affect workload.

Although it is a very minor change of 0.51 points, it is important that this value has not increased.

Taking again the references from Rebecca A. Grier's study [47], the results maintain similar low percentiles.

Weighted		**Raw/Unweighted**	
Overall	31,56	**Overall**	29,17
Diagnostic Subscores		Diagnostic Subscores	
Mental	116,67	Mental	33,33
Physical	46,67	Physical	20,00
Temporal	130,00	Temporal	33,33
Performance	100,00	Performance	30,00
Effort	73,33	Effort	26,67
Frustration	20,00	Frustration	13,33

Table 6: NASA-TLX results for the second Experiment

Regarding the components that have a greater impact on the total load, it is worth noting the loss of importance of frustration.

5.2.4 Conclusions

The results are really successful. The procedure has not only proven to be effective but to be simple, unobtrusive, effective and viable. The objectives set are taken for granted with this proposal.

Still it is necessary to point out the persistence of the possible pain point related to granting permissions to share screen.

Despite the efficiency of the procedure, it is also possible to evaluate in the future if using the internal zoom chat to access Maze does not make the procedure more complex.

5.3 Final Version

After validating the effectiveness and validity of the procedure, it is finally possible to make a proposal for the remote usability testing procedure.

5.3.1 Proposal

In order to carry out the procedure properly, the moderator must have the following:

- Informed Consent and Legal Documents. Coded in the **application the team wants**.
- **Zoom** room
- **Maze** test with the integrated Prototype. This prototype should be imported from **Adobe XD**, **Figma**, **inVision**, **Marvel** or **Sketch**.
- Forms (if necessary). Coded in the **application the team wants**.
- Other possible elements that the team considers necessary for its usability test.

Figure 45: Phases of the Remote Usability Testing Procedure

The procedure is defined in the following steps.

1. **Invitation:** The user is summoned to a meeting on the agreed date. For this, a standardized message is sent to the user by email. This message should include:

   - Zoom call link and access instructions.
   - Link to informed consent or other legal documents. The tool to use is at the discretion of the team.
   - The devices the user will need. Presumably a computer or tablet and a mobile.
   - Soft request to download the Zoom app. The user will also be given the opportunity to download it during the call if preferred.
   - It must be announced that during the procedure the user is going to share their mobile screen.

2. **Introduction:** Call starts in Zoom.

   The user is asked to confirm that they have signed the informed consent and the pertinent legal documents. If the user has not done it, the moderator will give instructions and support to do so.

   Once the documents have been signed and confirmed, the session can begin to be recorded.

   Next, the project that is being carried out is explained in detail. The process to be followed and its objectives are also explained.

   If there are questions or forms to fill out prior to prototype run, they are done at this time.

3. **Setting up the mobile:** A new standardized message is sent through email with the Zoom call link and access instructions. This link can also be open by the user from the meeting invitation message.

If the user does not have Zoom on his mobile or decided to not download it, the interview give him instructions to download it. Opening the access link, it will take the user to the corresponding store page to start the download. The download can be done also by searching for Zoom from the store.

Once Zoom is available, the user can join the call from their mobile using the access link or the credentials sent.

Once the mobile phone joins the call, the user is guided to share their mobile screen by clicking on the big green Share button and then selecting Screen.

4. **Remote Usability Test:** A new standardized message is sent through email with the Maze link .

   When the user open this link, the browser will open showing the test created.

   The moderator will guide the user through the Maze test. This process will change completely based on the needs and reality of the prototype to be evaluated.

   It is important to indicate to the user the importance and necessity of executing the think aloud protocol.

   Once the test in Maze is finished, the moderator can remove the user's mobile phone from the call.

   Next, if a questionnaire is to be carried out, the moderator will share their computer screen, showing the form to be completed. In case the user needs support for an additional file, a new standardized message is sent through the communication channel established with the needed files.

   The moderator will guide the user through the forms, being the one who completes the questions following the user's instructions.

   The discussion and the questions that the moderator want to ask can be done while the user continue on the call.

5. **End and farewell:** Once the usability test is finished and it is considered that all the information has been obtained, the moderator can say goodbye to the user.

   If the moderator wishes, the recording can be stopped. In any case, the recording will stop automatically when the moderator close the call.

   The procedure ends here and the moderator can close the call. Next, the call recording will be processed to generate the files corresponding to the meeting.

6 Discussion

This section presents an analysis of the status of the remote usability testing procedure. It includes a review of both the pain points that remain and possible alternatives to the conduct of the procedure. A number of possible lines of research to further improve the remote usability testing procedure are also presented.

6.1 Pain points

Despite trying to streamline the procedure as much as possible, there are still some potential pain points. The persistence of these is due to the limitations of the software or techniques available.

It would be advisable to continue investigating in the future how to alleviate these points in order to further optimize this procedure.

In any case, it is needed to be clear about these points in order to know how to react to them during the execution of the procedure.

Share screen permissions: The main possible critical point for the development of the procedure is when it comes to screen sharing in Zoom. If it has not been done previously, the mobile will request to provide permissions to be able to share screen.

This request may vary on each device, being confusing on some mobiles. It can help to inform the user of this and give him support. The camera of the computer or tablet can help in cases of particular confusion.

Internet connection: It is a good idea to be sure that the user has a decent internet connection. Although the procedure does not demand a large amount of available data, a very poor connection can make the procedure impossible.

Call setup: Users generally had no complications entering the call. But, is problems arise, there is no simple way to provide support in this phase of the project.

Memory Space: The user must download Zoom on their mobile. The amount of memory required for this is very small, but there is still the possibility of eventually running into problems.

This problem should be made even more sporadic by knowing in advance that users should download Zoom. If you do, they can fix the problem for you. Still, if they decide to do it during the call, this small possibility still exists.

6.2 Possible Alternatives

Despite the proven efficiency of the procedure, it is possible to emulate it with certain changes without greatly affecting it.

These proposals seek to offer some additional flexibility to adapt the procedure to the needs of different projects.

Forms: Despite the proposal of the procedure to fill in forms, there are alternatives.

It is possible to opt for a more traditional system and request the answers simply during the conversation. Screen sharing with questions can help users understand the questions. It also allows to save data directly without the need to transcribe it. However, including these questions directly in the conversation can make the discussion more fluid.

Another alternative is to integrate these forms within Maze. This option also allows to store the data easily. In addition, since the screen is being shared, the moderator will be able to see the user and give him support or help if it's needed.

This option can be very positive, but it surely requires having premium access to Maze. This is the main reason to remove it from the main procedure.

If this option is taken, if the answers are to be completed from the mobile, it is recommended to avoid long answer questions.

Prototype: The procedure is based on using Maze to share the prototype. If the prototype is in advanced stages of development, you may have a native application to use.

In this case, it would be important to pay special attention to how to share this application. It is recommended to do this once screen sharing is in progress. In this way, it is much easier to assist the user by being able to follow her movements.

Calling App: The call could be made from an application other than Zoom if decided.

The main advantage of Zoom is the no need to register or have an account. However, if desired, you could use one of the other applications that allow screen sharing.

This may be of interest if the users to be evaluated already belong to an internal team such as Teams or Slack.

In any case, these solutions need to download to the mobile so the procedure should not change much.

Retrospective Thinking-aloud: The procedure could be adapted to perform RTA if deemed valuable. To perform this technique, a review phase of the interaction would have to be added after completing the tasks in Maze.

With the current formulation, it would be a little more complicated. The interaction is currently being recorded during the call recording by Zoom. The problem is that this Zoom recording is not stored until the call ends. Therefore, it would be necessary to close the call, wait for the recording to process, and restart the call again. Through the screen sharing option, the previous call could be viewed.

As a possible positive point, this could help to solve problems with the time limitation of the free version of Zoom. This is because the realization of this second call would no longer need to integrate the user's mobile. As it is not a 3-person call, the time limitation would not exist.

Despite this, the process would be not clear and slow. This is why it would be best to use an alternative recording tool. After the interaction is over, the recording

could be stopped and processed. During this time, the user could be asked questions so as not to leave him on hold.

This alternative technique should not be incompatible with Zoom recording. Although the interaction was also recorded by another application, it would be useful to record the entire process as well. This is especially important if you ask questions while the interaction recording is processing.

6.3 Research continuation

Despite the positive results of the procedure evaluation, it would be possible to continue experimenting to further improve the process.

The procedure could be improved in three different aspects. These in turn would need research for their proposal and its validation.

Focusing future research in any of these directions would help improve the usefulness and quality of the procedure.

Of course, the investigation would not have to be restricted to just one of these. The joint study of them could be, even more, to improve the proposed procedure.

Streamline the procedure: There is room for improvement in terms of the simplicity of the procedure. The mentioned pain points can still try to alleviate.

No optimal way has been found to treat these points with current tools. However, it would be possible to try to find alternatives to the tools used.

If alternatives to the proposed tools are found, it would be necessary to carry out a new experiment to check the opinion of users regarding pain points. Additionally, it would be necessary to check if the information obtained is not less valuable than with the current procedure.

Improve the effectiveness of the procedure: Having ensured the simplicity of the procedure, it may be helpful to make it more useful. Tools and phases that allow better data collection could be investigated.

Although the procedure has proven to be effective in collecting usability test data, this can be improved.

The usability test has been defined as an external element to the procedure. This is because it must adapt to the reality of the prototype to be used and the needs of the project. However, it is possible that the procedure covers this part in a more direct way.

If it were decided to take this approach, it would be necessary to evaluate the information obtained in comparison with previous iterations. However, it would also be important to try to maintain the initial simplicity goal.

A further interesting study could be to find a balance between the possible additional difficulty of the procedure as opposed to the greater amount of information collected.

Make the procedure more flexible: In this work, some possible alternatives to the current procedure have already been offered. However, it is possible to find more.

A more flexible procedure would facilitate its use in even more diverse case studies. Not only would it be more useful for more people, but it would also allow the same team to have the same procedure for very different cases.

It might also be of interest to further integrate this flexibility into the proposal. Alternatives to the procedure could be included in the procedure formulation instead of an annex to it.

7 Conclusions

The proposal of the procedure is considered valid with respect to the defined objectives.

Despite the growing popularity of remote testing, there is not much standardization on how to do it using mobile applications. In fact, there are only a few existing tools oriented towards this process.

However, after an analysis of the current literature and the available applications, it has been possible to propose an efficient procedure to conduct tests remotely with mobile applications.

This procedure is simple to drive and follow. Plus, it's as non-intrusive as possible for users, with hardly any downloads or registrations required.

Additionally, the results obtained are valid enough to make it worthwhile.

This has been possible after two confirmatory experiments. The first has made it possible to verify the effectiveness of the procedure. It has also made it possible to detect certain pain points that should be solved.

After the second proposal and validation of the procedure, most of these pain points have been solved. Despite this, certain points must be taken into account to help carry out the procedure.

Finally, apart from these pain points, a series of possible alternatives are offered. These seek to make the procedure more flexible to adapt to different realities in different case studies.

7.1 Research Questions

A series of research questions were posed at the beginning of the project. These have guided the conduct of the study and allowed to answer them.

1. **How to perform remote usability testing for mobile payment applications?:** Following the phases established in the proposed procedure and using the proposed tools, it is possible to perform remote usability testing for mobile payment applications.

   In fact, the proposed procedure aims to be functional with different types of mobiles application.

   This procedure makes the process easy for both the research team and the users interviewed. It also ensures that valid information is obtained to guarantee the usefulness of this kind of usability testing.

2. **What tools can be used to perform remote usability testing for mobile payment applications?:** The use of email, Zoom, Maze (and a compatible prototyping application) and a forms application such as Google Forms allows for the realization of remote usability testing for mobile payment applications.

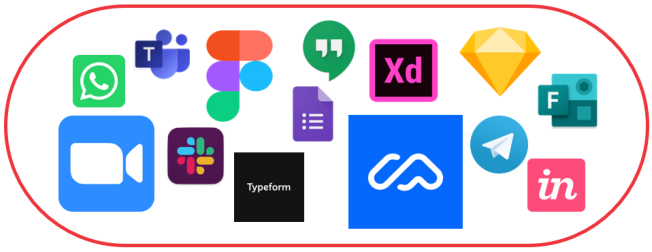

Figure 46: Proposed applications to be used in Remote Usability Testing with mobiles

The procedure presented offers the tools, resources and steps necessary to easily replicate the process. Not only that, but the study also presents a series of alternative tools for conducting remote usability testing.

It would have been useful to have been able to include a larger number of applications to answer this question. However, the lack of software dedicated to remote usability testing has not allowed to have more options to analyze.

This is especially true for mobile testing as well. Maze, for example, has been a very powerful and necessary tool, but a more valuable answer to this question could have been given if more applications were available.

It is also positive to have more tools to use if there is easy integration and communication between them.

3. **How can the remote usability testing procedure be adapted to different needs?:** The proposed procedure is flexible and allows the use of different applications (such as Microsoft Teams or Slack), different techniques (such as RTA) and adapts to different phases of development (allowing the use of a native application instead of a prototype).

   The procedure offers comprehensive tools to ensure the success of remote usability testing such as the use of Zoom or Maze. Even so, it tries to be flexible and not restrict the use of other possible alternatives. An example of this is that selecting Maze allows to choose between different prototyping tools such as Figma or Sketch.

   In addition, a series of proposals and alternatives have been presented. These seek to make the procedure more flexible in order to adapt to different needs.

   These proposals not only offer the use of alternative tools but also cover changes in the methodology of remote usability testing.

   A wider offer of dedicated software and tools would have allowed a more complete answer to this research question similar to the previous one.

   Nevertheless, the proposals presented offer a sufficient range of margin and flexibility to be able to positively answer this research question.

Having answered the research questions posed, it is possible to consider the study and the project a success.

7.2 Software improvements

As mentioned, the use of remote usability testing with mobile applications continues to grow, specially with payment applications. This trend could continue after the successful results in proposing and conducting the remote usability testing procedure.

This growing need for remote usability testing with mobile applications is a great opportunity to create new software.

Maze has proven to be a perfect tool due to its simplicity. It is simple to configure, it is versatile and avoids the user hassles, downloads and installations. However, much more could be done.

It could be positive to have applications that allow the execution of this type of usability test, also offering an environment for communication. Also being able to record the camera and avoid the need to open more applications would make a process like the one proposed much easier.

It would also facilitate the process if you could participate in video calls on your mobile from a browser. Although almost all the solutions analyzed allowed making calls from the browser on a computer, none allowed it on a mobile phone. Avoiding downloads not only makes design easier, but helps avoid scaring off the users.

Since the biggest pain point is giving Zoom permissions, it might help if mobile software developers were aware of this. Although it is a one-time operation, we have almost accidentally detected this problem several times.

The solution of this problem is in principle beyond our reach. However, if all the mobiles had had a simple system to give permissions, this problem would not have arisen.

7.3 Impact

The existence of a standardized procedure for conducting remote usability testing with mobile applications can be very useful. Having a predefined procedure can help many research teams to perform this type of usability testing and benefit from the advantages of remote testing.

The standardization of software development process has proven also positive effects by itself. These effects are expected to apply also to the procedure presented in this paper. The advantages include the improvement in project performance and the software quality [61]. It can also lead to a boost in employee morale as they can easily master the process [62].

It would also be positive to raise collective awareness about the importance of conducting this type of testing with mobile devices. It is well known that mobile applications have been on the rise for years and yet the attention paid to testing these applications remotely has not increase at the same speed. This is plausible in payment applications but it can be extended to all types of mobile applications.

Thus, this study and the proposed procedure seeks not only to offer a help for research teams but also to give a wake-up call. The research, user experience and software development communities should be made aware of the need to conduct remote usability tests with mobile phones.

Since standardizing software development processes can also lead to a reduction in maintenance costs [63], it would be positive if the companies invest part of the extra profits in improving the available tools.

If the impact of this work is scaled up, it would be of great interest to continue research on how to make the presented procedure more flexible and optimized. It would be very positive for the procedure to benefit from the experience of different research teams and different case studies.

Hopefully in the not too distant future this practice will become sufficiently established to have more standardized guidelines supported by large companies and teams. And I hope studies like this one can help advance in that direction.

References

[1] C. Tam and T. Oliveira, "Literature review of mobile banking and individual performance," *International Journal of Bank Marketing*, vol. 35, no. 7, pp. 1042–1065, 2017.

[2] A. Pal, T. Herath, R. De', and H. Raghav Rao, "Why do people use mobile payment technologies and why would they continue? an examination and implications from india," *Research Policy*, vol. 50, no. 6, 2021.

[3] Reuters., "The next decade in digital payments," 2018.

[4] A. Meola, "Overview of the fintech industry: stats, trends, and companies in the ecosystem market research report," 2021.

[5] S. Duarte, M. Costa, M. Brito, A. Miranda, and M. Au-Yong-Oliveira, "Bank and online payment apps : How beneficial to society in the xxi century," in *Iberian Conference on Information Systems and Technologies, CISTI*, vol. 2020-June, 2020.

[6] Usability.gov, "Usability testing," Nov 2013.

[7] J. Sauer, A. Muenzberg, L. Siewert, A. Hein, and N. Roesch, *Remote Testing of Usability in Medical Apps*, vol. 320 LNICST of *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*. 2020.

[8] R. J., *Learning About the User Experience on the Web With the Phone Usability Method*. Human Factors and Web Development, 2002.

[9] J. R. Hill, A. B. Harrington, P. Adeoye, N. L. Campbell, and R. J. Holden, "Going remote-demonstration and evaluation of remote technology delivery and usability assessment with older adults: Survey study," *JMIR mHealth and uHealth*, vol. 9, no. 3, 2021.

[10] A. J. Bernheim Brush, M. Ames, and J. Davis, "A comparison of synchronous remote and local usability studies for an expert interface," in *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1179–1182, 2004.

[11] A. C. McLaughlin, P. R. DeLucia, F. A. Drews, M. Vaughn-Cooke, A. Kumar, R. R. Nesbitt, and K. Cluff, "Evaluating medical devices remotely: Current methods and potential innovations," *Human factors*, vol. 62, no. 7, pp. 1041–1060, 2020.

[12] J. Cuadros, V. Serrano, J. Garcia-Zubia, and U. Hernandez-Jayo, "Design and evaluation of a user experience questionnaire for remote labs," *IEEE Access*, vol. 9, pp. 50222–50230, 2021.

[13] S. L. Henry and L. McGee, "Accessibility - w3c."

[14] N. Mallat, "Exploring consumer adoption of mobile payments - a qualitative study," *Journal of Strategic Information Systems*, vol. 16, no. 4, pp. 413–432, 2007.

[15] S. G. Dewan and L. da Chen, "Mobile payment adoption in the us: A cross-industry, crossplatform solution," *Journal of Information Privacy and Security*, vol. 1, no. 2, pp. 4–28, 2005.

[16] K. Pousttchi, "A modeling approach and reference models for the analysis of mobile payment use cases," *Electronic Commerce Research and Applications*, vol. 7, no. 2, pp. 182–201, 2008.

[17] "Mobile payment adoption: Classification and review of the extant literature," *The Marketing Review*, vol. 13, no. 2, pp. 167–194, 2013.

[18] S. Teng and K. W. Khong, "Examining actual consumer usage of e-wallet: A case study of big data analytics," *Computers in Human Behavior*, vol. 121, 2021.

[19] B. Burrell-Squires, "Mobile money vs mobile banking – what's the difference?," 2020.

[20] "Difference between mobile money and mobile banking (with table)."

[21] I. R. de Luna, F. Liébana-Cabanillas, J. Sánchez-Fernández, and F. Muñoz-Leiva, "Mobile payment is not all the same: The adoption of mobile payment systems depending on the technology applied," *Technological Forecasting and Social Change*, vol. 146, pp. 931–944, 2019.

[22] E. Valcourt, J.-M. Robert, and F. Beaulieu, "Investigating mobile payment: supporting technologies, methods, and use," in *WiMob'2005), IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, 2005.*, vol. 4, pp. 29–36 Vol. 4, 2005.

[23] F. Liébana-Cabanillas, I. R. de Luna, and F. Montoro-Ríosa, "Intention to use new mobile payment systems: A comparative analysis of sms and nfc payments," *Economic Research-Ekonomska Istrazivanja*, vol. 30, no. 1, 2017.

[24] "Download limit exceeded."

[25] "Report says that sms is not ideal for emergency communications."

[26] D. Wave, "Qr code, iso/eic18004," 2000.

[27] "Essentials for successful nfc mobile ecosystems," 2008.

[28] X. Chen, K. Choi, and K. Chae, "A secure and efficient key authentication using bilinear pairing for nfc mobile payment service," *Wireless Personal Communications*, vol. 97, no. 1, 2017.

[29] K. Grassie, "Easy handling and security make nfc a success," *Card Technology Today*, vol. 19, no. 10, pp. 12–13, 2007.

[30] G. Hancke, "Eavesdropping attacks on high-frequency rfid tokens," 2008.

[31] C. Otto, "A la desesperada: 27 bancos españoles lanzan una 'fintech' para luchar con internet," Aug 2016.

[32] S. Vergara, "Bizum como método de pago en tu tienda online woocommerce," Nov 2020.

[33] N. Sarriés and A. Stumpf, "La banca baraja bizum como solución paneuropea de pagos," Feb 2020.

[34] "Bizum alcanza los 10 millones de usuarios," Aug 2020.

[35] Efe, "Bizum llega a 15 millones de usuarios y mueve ya más de 18.000 millones," Mar 2021.

[36] "Human-centred design for interactive systems, iso 9241-210," 2019.

[37] S. Riihiaho, *Experiences with usability testing: Effects of thinking aloud and moderator presence*. PhD thesis, Aalto University, 2015.

[38] C. Barnum, *Usability Testing Essentials*. Usability Testing Essentials, 2011.

[39] J. S.Dumas and J. Redish, *A Practical Guide to Usability Testing*. Intellect Books, 1999, 1993.

[40] J. Romano Bergstrom and E. Olmsted-Hawala, "Think-aloud protocols: Does age make a difference?," 01 2012.

[41] S. Dray and D. Siegel, "Remote possibilities? international usability testing at a distance," *Interactions*, vol. 11, pp. 10–17, 01 2004.

[42] J. Brooke, "Sus: A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, 11 1995.

[43] M. Del Rocio Sevilla-Gonzalez, L. M. Loaeza, L. S. Lazaro-Carrera, B. B. Ramirez, A. V. Rodríguez, M. L. Peralta-Pedrero, and P. Almeda-Valdes, "Spanish version of the system usability scale for the assessment of electronic tools: Development and validation," *JMIR Human Factors*, vol. 7, no. 4, 2020.

[44] A. Smyk, "The system usability scale how it's used in ux," Mar 2020.

[45] I. Arquer and C. Nogareda, "Estimación de la carga mental de trabajo: el método nasa tlx. ntp 544.," 1999.

[46] E. Díaz Ramiro, S. Rubio Valdehita, J. Martín Garcíaa, and L. Luceño Moreno, "Estudio psicométrico del Índice de carga mental nasa-tlx con una muestra de trabajadores españoles," *Revista de Psicología del Trabajo y de las Organizaciones*, vol. 26, pp. 191 – 199, 12 2010.

[47] R. A. Grier, "How high is high? a meta-analysis of nasa-tlx global workload scores," in *Proceedings of the Human Factors and Ergonomics Society*, vol. 2015-January, pp. 1727–1731, 2015.

[48] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '90, p. 249–256, Association for Computing Machinery, 1990.

[49] A. Rahdan, "Jakob nielsen's 10 heuristics for user interface design," Jun 2020.

[50] P. Sheldon, D. L. Giudice, and S. Compton, "Have you "signed in" with facebook recently?," Jul 2017.

[51] T. . Jordan, "2020 tools survey results."

[52] I. Sherr, "Discord, slack for gamers, tops 250 million registered users," May 2019.

[53] S. Salguero, "Todo lo que deberías saber sobre discord (el slack de los gamers)," Feb 2019.

[54] M. Guay, "The 9 best online form builder apps in 2020," Jun 2014.

[55] "Formas de pago online más usadas en españa," Jan 2021.

[56] R. A. Virzi, "Refining the test phase of usability evaluation: How many subjects is enough?," *Human factors*, vol. 34, no. 4, pp. 457–468, 1992.

[57] P. Fava, "Estas son las regiones de españa en las que mejor se come según la ciencia," Feb 2019.

[58] "Las tendencias que están transformando los hábitos de compra en el gran consumo," Feb 2020.

[59] Gobierno de España: Ministerio de Agricultura, Pesca y Alimentación, "Análisis de consumo alimentario."

[60] P. Barrachina, "Dama levante," Aug 2020.

[61] J. Y. . Liu, V. J. Chen, C. . Chan, and T. Lie, "The impact of software process standardization on software flexibility and project management performance: Control theory perspective," *Information and Software Technology*, 2008.

[62] B. Brandall, "Why process standardization improves quality, productivity, and morale: Process street: Checklist, workflow and sop software," Nov 2020.

[63] N. Contributor, "6 reasons standardizing your software is essential," Jul 2021.